SMALL MOLECULE RECOGNITION OF HOMOPURINE NUCLEIC ACID STRUCTURES

A Thesis
Presented to
The Academic Faculty

by

Özgül Persil Çetinkol

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology
August 2008

SMALL MOLECULE RECOGNITION OF HOMOPURINE NUCLEIC ACID STRUCTURES

Approved by:

Dr. Nicholas V. Hud, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Adegboyega K. Oyelere
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Donald F. Doyle
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Roger M. Wartell
School of Biology
Georgia Institute of Technology

Dr. Kirill S. Lobachev
School of Biology
Georgia Institute of Technology

Date Approved: June 26, 2008

To my mother Semiha Persil,

To my father Yunus Persil,

To my sisters Özge and Selin Persil

and

To my husband Mehmet Çetinkol

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my advisor, Professor Nicholas V. Hud. Your passion for science, academic sincerity, patience to teach, creativity, and perfectionist attitude have helped me grow both professionally and personally. You have been a remarkable teacher. I will remain forever grateful for being accepted to your,"our" ,lab, for your endless support and trust even when I didn't trust myself, for your positive attitude and for being always open to talk and understand.

Needles to say, I would like to thank to the exceptional former and recent members of the Hud Lab; Christine Conwell, Igor Vilfan, Swapan Jain, Pat Bishop, Cathy Santai, Tumpa Sarkar, Heather Bean, Eric Horowitz, Aaron Engelhart, Ragan Buckley and Seth Lilavivat thanks for creating such a fun friendly atmosphere and for your invaluable input and support. Tumpa and Heather, my fellow old ladies, I could have completed the program if I didn't have your shoulders to lean on. I would like to thank also to dearest Tatsuya "T" Maehigashi, Derrick Watkins and Srividya Mohan for opening the door for me whenever I needed to run down.

I also would like to thank to my committee members; Profs. Donald Doyle, Kirill Lobachev, Adegboyega K. Oyelere, Roger Wartell, and Loren Williams for their invaluable input to my graduate education. I would like to thank Dr. Janez Plavec for his invaluable insight on my initial studies and his encouragement about my scientific career. I would like to thank Dr. Les Gelbaum for his kindness, patience and guidance through my NMR experiments. I would like to thank Mr. David Bostwick for all his help with the mass spectroscopy. I would like to thank the members of Schuster lab, especially Dr. Joshy Joseph, for their help with the oligonucleotide synthesis. I also would like to thank

to all of the staff of the School of Chemistry and Biochemistry, especially Dr. Cam Tyson, and the staff of Institute of Bioengineering and Bioscience (IBB) for all their help and support in the last six years.

I would like to thank to my friends in Atlanta. Selma Bakbak, Gözde Güler, Alpay Kimyonok, Mehmet Kütükçü, Aslı Ovat, thanks for your unconditional friendship and support. I will always cherish our friendship.

The graduate school was only the continuation of my undergraduate education. Therefore, I would like to thank all my undergraduate professors, especially Prof. Mehmet S. Somer and Dr. Durata H. Ertek, for encouraging me for applying to graduate school. I also would like to thank to my lifelong friends Sema Akboğa, Eylem Akdoğan, Melike Akşit, Deniz Bolsoy Erdem, Şebnem Eşsiz, Nurdan Kabay, Pınar Tan and Yeliz Utku for making my life enjoyable. I would like to express my deepest gratitude to Dr. Can Paker and Türk Henkel Company for their support to pursue my education. I will be grateful forever.

Finally, I would like to thank my family to whom this thesis is dedicated. Words alone will not be enough to express my love and appreciation for you. Thanks for your endless unconditional support and love. Mom and Dad thanks for being so wonderful, thanks for letting me choose my own dreams and encouraging me to pursue them. Thanks for teaching me the importance of education, and for always believing in me. My beautiful sisters, you will always be the sweetest troublemakers and best friends I have and I can't thank you enough. Mehmet; you mean the world to me. I wouldn't have been here without your love and support. Thanks for being always here and holding me. "And nothing else matters".

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS AND ABBREVIATIONS

DNA	deoxyribonucleic acid
RNA	ribonucleic acid
bp	base pair
A	adenine
T	thymine
G	guanine
C	cytosine
U	uracil
I	inosine
D	diaminopurine
X	xanthine
isoG	isoguanine
7-deaza	7-deaza deoxy adenosine
UV	ultraviolet
CD	circular dichroism
NMR	nuclear magnetic resonance
λ	wavelength (nm)
θ	ellipticity
$\Delta\varepsilon$	molar circular dichroism
mdeg	millidegrees (ellipticity units)
T_m	melting temperature

ΔG	Gibbs free energy
ΔH	change in enthalpy
ΔS	change in entropy
R	universal Gas Constant
I	incident light
I_0	transmitted light
rms	root mean square
PAGE	polyacrylamide gel eletrophoresis
ESI	electrospray ionization
MS	mass spectrometry
HPLC	high performance liquid chromatography
TLC	thin layer chromatography
AFM	atomic force microscopy
M/Z	mass to charge ratio
E_a	activation energy
K_d	equilibrium dissociation constant
Ka	equilibrium association constant
H-bond	hydrogen bond
TEAA	triethyl ammonium acetate
CMC	critical micelle concentration

SUMMARY

Small molecules that bind to nucleic acids have attracted extensive interest for more than 40 years, largely due to the activity of some such molecules as therapeutic drugs or as mutagenic agents. The role of DNA and RNA in many cellular functions immediately represents an opportunity for the treatment of diseases of genetic origin by the targeting of specific nucleic acid structures with small molecules. However, we are only beginning to understand the interactions of small molecules with nucleic acids at a level of detail sufficient to allow the design of molecules that bind tightly and selectively to specific nucleic acid structures. For example, the binding of small molecules to non-canonical nucleic acid structures (i.e. non-B-form and non-A-form structures), such as triplex and quadruplex DNA, has been a major focus of rational drug design. While these efforts have begun to show success, we are still far from the point at which small molecules can be designed to selectively bind one particular nucleic acid structure.

In addition to their biological and medicinal importance, nucleic acids have emerged as the material of choice for the creation of supramolecular assemblies in nanotechnology. Small molecule-DNA interactions therefore represent another dimension for expanding the functionality and dynamics of DNA-based nanotechnology. In regards to small-molecule-driven nucleic acid assembly, the binding of small molecules to the nucleic acid bases through non-covalent interactions has been proposed by our laboratory to have even played a central role in the origin of life by facilitating nucleic acid assembly via stacking interactions.

The thesis topic entitled above involves the use of small molecules as a general means to drive nucleic acid assembly and structural transitions. A primary goal has been

to gain a fundamental understanding of the physical and chemical principles that govern small molecule intercalation of nucleic acids by using spectroscopic techniques.

We have shown that coralyne, a crescent-shaped small molecule, can assemble homo-adenine DNA and RNA sequences into anti-parallel duplexes at neutral pH, a structure containing putative purine-purine (A·A) base pairs that is otherwise unstable. The importance of the structure of the small molecule in the recognition and stabilization of A·A base pairing has been established by experimental evidence. We further provide structural evidence for the putative A·A base pairing, that is stabilized by coralyne and molecules of similar size and shape. Our hypothesis that *planar molecules that are slightly too large to intercalate Watson-Crick base pairs might intercalate the larger purine-purine base pairs* has led to the design of a new class of small molecules that tightly bind purine-purine duplexes with excellent selectively. We have demonstrated that azacyanines can exhibit strong and selective association with a human telomeric sequence that forms a unimolecular G-quadruplex in solution. The synthetic accessibility of azacyanines makes the class amenable to library preparation for high-throughput screening.

Together, the findings reported in this thesis provide further evidence to the robust and versatile nature of selective small molecule recognition of nucleic acids, especially purine-purine duplexes.

CHAPTER 1

INTRODUCTION

1.1. DNA STRUCTURE

All living organisms depend on nucleic acids for their survival and procreation. Nucleic acids are responsible for the storage, conservation, expression and transmission of the genetic information. Structurally speaking, nucleic acids are linear polymers built from monomeric units called nucleotides. A nucleotide has three components: sugar, heterocycle and phosphate. The sugar, ribose or deoxyribose, is connected by the β-glycosyl linkage to one of the heterocyclic bases (purine: adenine, guanine or pyrimidine: cytosine, thymine, uracil) to produce nucleosides. If a nucleoside is phosphorylated at the 3' or 5' end, then it is called a nucleotide. The structure of the bases and the nucleotides are given in Figure 1.1.

For the first time, in 1944 Avery et. al. demonstrated the role of DNA as the carrier of our genetic information [1]. In 1949 Erwin Chargaff et. al. demonstrated that the amount of adenine is always equal the amount of thymine. The same rule is also valid for the guanine and cytosine, even though different organisms have different amounts of DNA and different ratios of adenine to guanine [2, 3]. A true scientific breakthrough was made in 1953 when James Watson and Francis Crick suggested a helical structure for DNA in their Nature article "Molecular Structure of Nucleic Acids" [4].

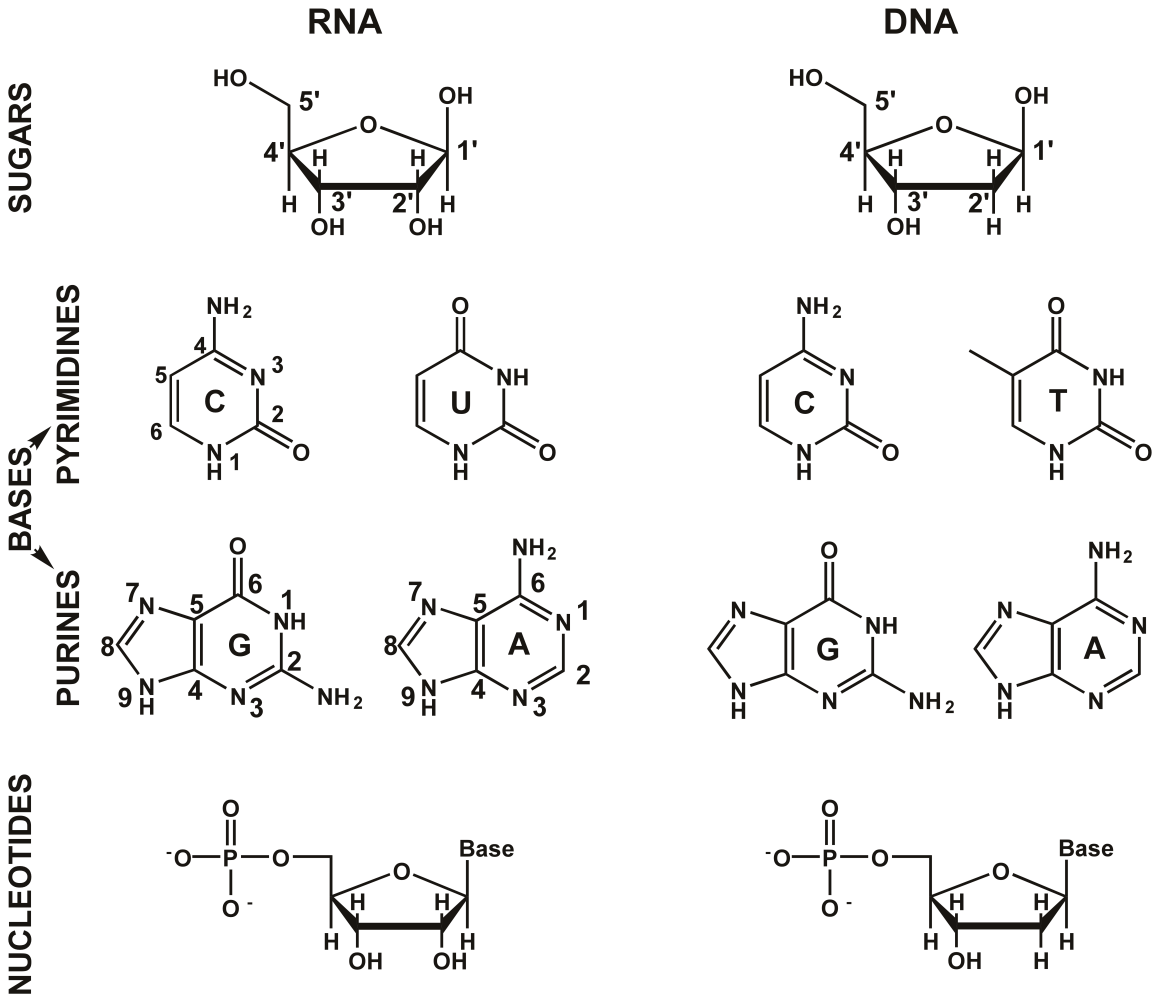

Figure 1.1: Structures of ribose and deoxyribose (top), bases (middle) and the nucleotide structure.

The Watson-Crick structure of so-called B-form DNA consisted of two polynucleotide strands that wind around each other with a right-handed twist to form the double helix. The strands are anti-parallel, the bases occupied the core of the helix and the sugar-phosphate chains are located on the outside. Hydrogen bonding between the bases from the opposite strand and the stacking of the bases on top of each other held the

helix intact. For the first time, Watson-Crick demonstrated that there are only two types of base pairs in genomic DNA; adenine-thymine and guanine-cytosine. The structures of "Watson-Crick base pairs" A·T and G·C are given in Figure 1.2.

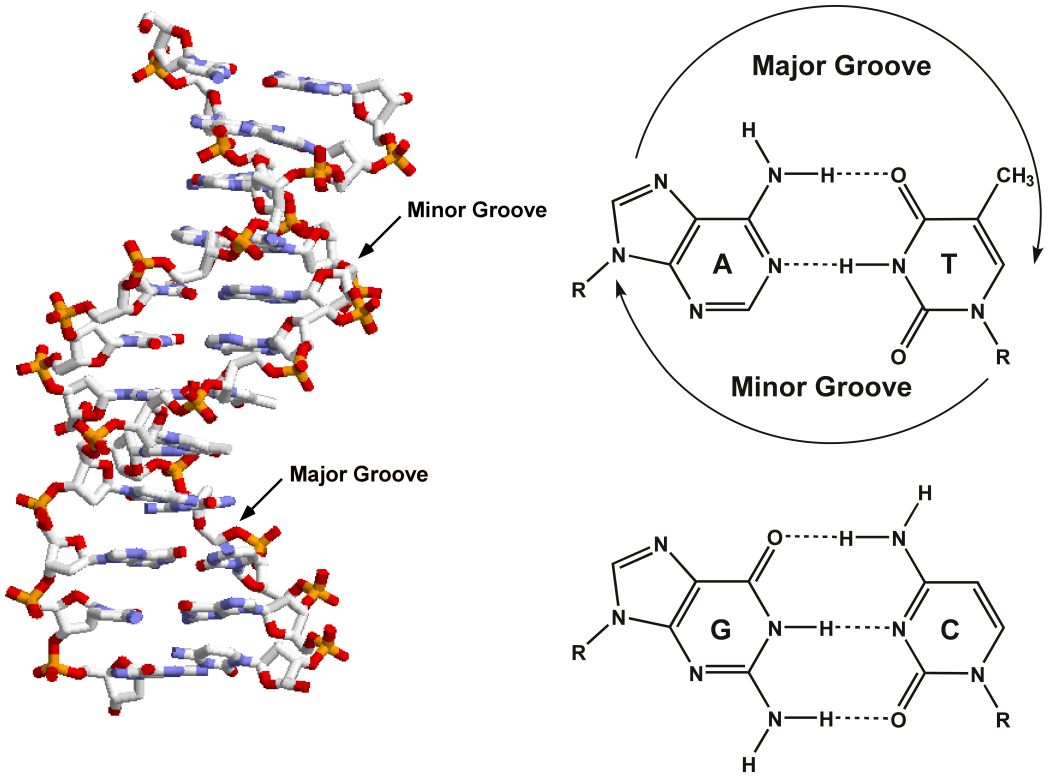

Figure 1.2: B-form DNA (left). The picture was created from the PDB id 1fq2 [5] with the Rastop program. Chemical structures of A·T (top right) and G·C (bottom right) Watson-Crick base pairing. The deoxyribose attached to the bases is represented by R.

Today, B-form DNA is regarded as the native form that is found in the cells of eukaryotes and bacteria under physiological conditions [6]. The two other commonly discussed double helical conformations of DNA are A-form and Z-form. A-form DNA is known as the dehydrated helical form of DNA that is obtained upon the conformational change of B-DNA when the humidity is reduced to 75% relative humidity. In contrast, RNA adopts the A-form conformation up to 100% relative humidity. Z-DNA is left-

handed and formed only by sequences that have alternating pyrimidine-purine sequences, and particularly alternating C and G bases. The major structural features for these three common conformations of DNA are given in Table 1.1

Table 1.1: Structural features of ideal A-, B- and Z-DNA. Table is adapted from Voet et. al [6].

Parameter	A-DNA	B-DNA	Z-DNA
Helical sense	Right handed	Right handed	Left handed
Diameter	~26 Å	~20 Å	~18 Å
Base pair per helical turn	11.6	10	12(6dimers)
Helical twist per base pair	31^o	36^o	9^o for pyr-pur steps $51^o$for pur-pyr steps
Helix pitch (rise per turn)	34 Å	34 Å	44 Å
Helix rise per base pair	2.9 Å	3.4 Å	7.4 Å per dimer
Base tilt normal to the helix axis	20^o	6^o	7^o
Major groove	Narrow and deep	Wide and deep	Flat
Minor groove	Wide and shallow	Narrow and deep	Narrow and deep
Sugar Pucker	C3'-endo	C2'-endo	C2'-endo for pyr C3'-endo for pur
Glycosidic bond	Anti	Anti	Anti for pyr Syn for pur

1.2. HIGHER ORDER STRUCTURES AND NON-CANONICAL BASE PAIRING

Under physiological conditions DNA is generally found as a double helix, however, it can also adopt higher order structures such as triplex or quadruplex structures. On the other hand, RNA is generally single stranded, but folds on itself to form secondary and tertiary structures that include helices, bulges, pseudoknots and hairpins. The higher order structures often involve different possible base pairing arrangements called “non-canonical base pairs” in addition to A·T and G·C Watson-Crick base pairs. With the determination of high resolution nucleic acid structures, the important role of the non-canonical base pairs in tertiary structures and function is now well established, particularly in the case of globular RNA structures [7].

Triplexes are generally formed by the placement of a third strand in the major groove of a DNA or RNA double helix. The interaction of the third strand, which can be either a pyrmidine or a purine strand, with the Watson-Crick duplex is through so-called “Hoogsteen” base pairing [7]. As a consequence, the central strand of the triplex needs to be a homopurine sequence to provide the necessary sites for the formation of both Watson-Crick and Hoogsteen base pairing, as shown in Figure 1.3. The interest in triplex structures increased in the last two decades mainly due to the idea of anti-gene therapy which is the idea of blocking the transcription of certain genes by the formation of a triplex structure between an exogenous oligonucleotide with a double stranded DNA sequence within the target gene [7].

Quadruplexes, which are four stranded DNA or RNA structures, are formed by G-rich sequences. In these structures, four guanines bases form G-quartets (also known as

G-tetrads) that stack on top of each other to form G-quadruplex structures. These G-quartets are stabilized by Hoogsteen base pairing, in addition to the coordination of closely spaced carbonyl groups by cations (Figure 1.4). The possible biological activity of G-quadruplexes suggested by the presence of G-rich sequences in telomere sequences and gene promoter regions have made G-quadruplex structures an active area of research in the last decade [8-12].

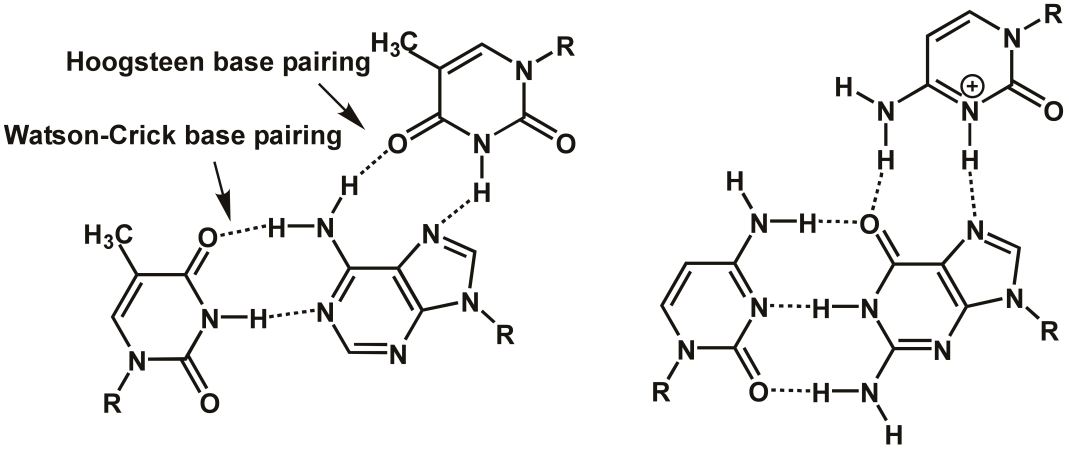

Figure 1.3: The most common triplex structures: T·A·T (left) and C·G·C$^+$ (right). R represents the sugar moiety attached to the base.

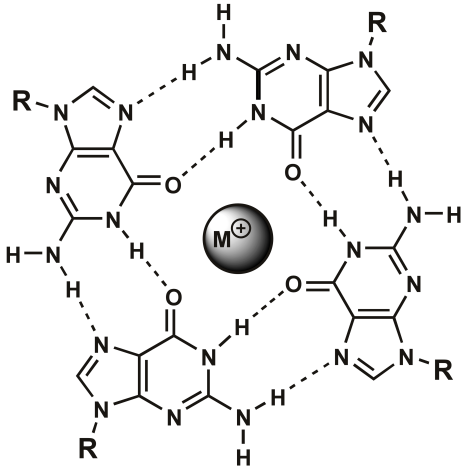

Figure 1.4: G-quartet structure. M represents the metal cation and R represents the sugar moiety attached to the base.

In addition to the triplex and quadruplex structures, RNA also adopts a variety of secondary and tertiary structures that include duplex regions. The most common of these secondary/tertiary structures are given in Figure 1.5 [7].

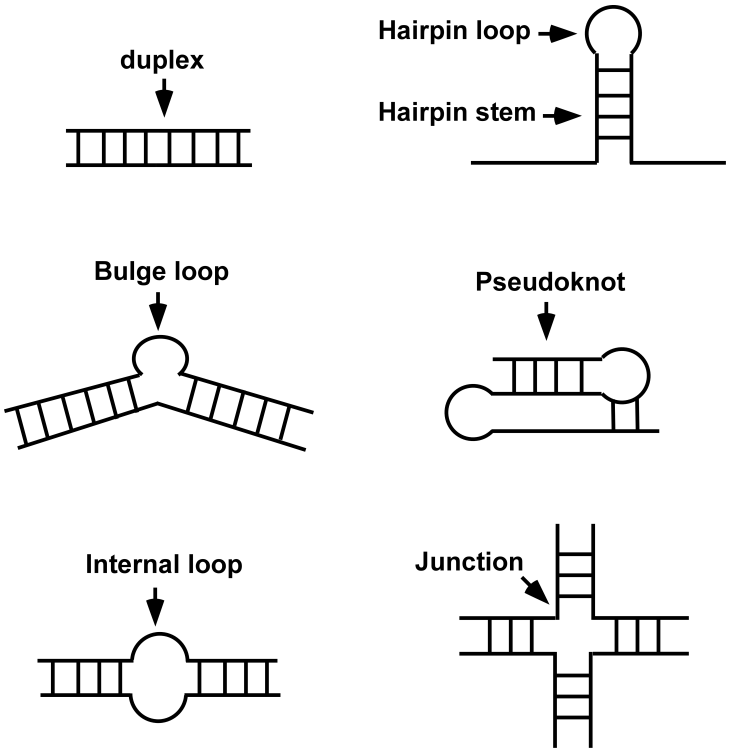

Figure 1.5: Common RNA secondary structures.

The most common non-canonical base pairs found in RNA structures are the G·U wobble, the trans A·G Hoogsten, the trans A·U Hoogsteen and the trans A·A Hoogsteen (Figure 1.6) [13]. The role of non-canonical base pairs in stabilizing these structures in biologically important RNAs, such as in the ribosome, natural ribozymes, and tRNAs has been well established mainly by structural studies [14, 15]. For example, most of the natural ribozymes, such as group I and group II introns and Hepatitis Delta Virus (HDV) ribozyme, have conserved G·U wobble pairs, as substitution of G·U wobble pairs with

other base pairs abolishes RNA function. As a particular example, the G1·U37 wobble pair at the P1 helix of HDV ribozyme is critical for its cleavage reaction [16].

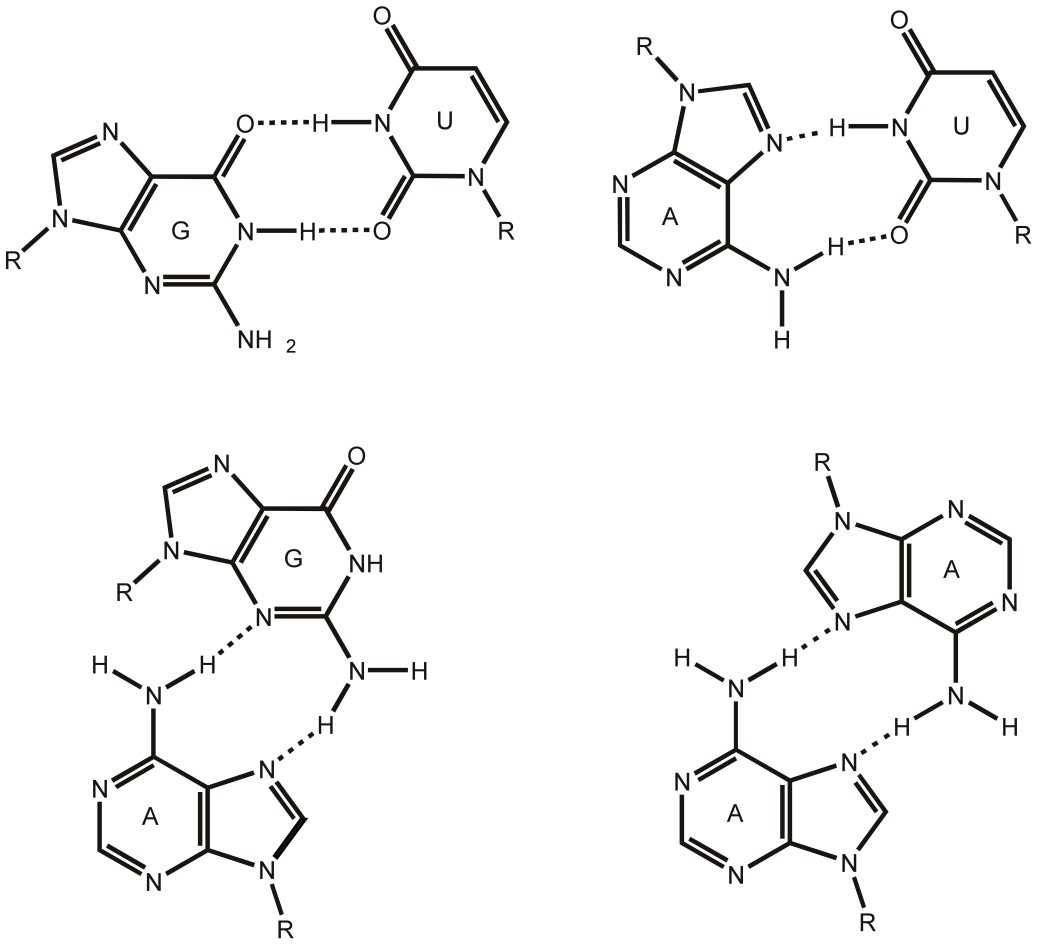

Figure 1.6: The most common non-canonical base pairs in RNA structures [13] .

1.3. SMALL MOLECULE NUCLEIC ACID INTERACTIONS

The binding of small molecules to nucleic acids has attracted extensive interest for more than 40 years, largely due to the activity of these molecules as therapeutic drugs or as mutagenic agents [17]. The role of DNA in cellular functions immediately represents an opportunity to treat diseases due of genetic origin by targeting specific DNA sequences with small molecules [18, 19]. The principal objective of targeting DNA

by small molecule binding in treating genetic diseases is to down regulate or abolish the function of certain genes. The first compounds discovered to bind to DNA were the sulfur mustards, which are highly toxic to all cells [20]. With the advances made since that time across several scientific disciplines, today we know of many small molecules that bind to DNA and are clinically proven as therapeutic agents such as echinomycin, imidazoacridone, distamycin and quarfloxin [19, 21-27]. All of these molecules need to bind to DNA tightly but also selectively in order to eliminate toxicity and off target effects. The concept of designing compounds that will bind tightly and selectively to cellular targets is part of the field known as "rational drug design" [28]. Consequently, the polymorphic structure of nucleic acids, i.e. the different conformations that nucleic acids can adopt in solution, provides unique opportunities for the design of selective, tightly binding small molecules [29, 30]. Of course, the structure of a small molecule is ultimately responsible for selecting between different nucleic acid conformations (and other cellular molecules), the so-called "molecular recognition" of nucleic acids by small molecules. An excellent example of molecular recognition by a small molecule is given by the work of Chaires et. al [31]. These researchers reported that daunorubicin binds preferentially to right handed DNA while its (-) enantiomer WP900 preferentially binds to left handed DNA. Moreover, under the solution conditions that favor left handed-conformation (Z-form) of [poly(dGdC)], binding of daunorubicin converted left handed nucleic acid structure to a right handed intercalated structure. The opposite was true for (-) enantiomer that it converted right handed nucleic acid structure to a left handed intercalated structure.

1.3.1 Mode of Binding

Small molecules that bind to DNA through non-covalent interactions can be grouped into three general categories (Figure 1.7). The first category discovered is intercalation, which was reported by Lerman in 1961 [32]. Based on the changes in viscosity of the DNA solution upon binding of acridines, Lerman proposed that acridines bind to DNA such that the length of the DNA increases. Lateral, X-ray fiber diffraction studies confirmed his proposal that the planar, aromatic portion of these small molecules slide between adjacent base pairs without disturbing the overall stacking interactions (i.e. the adjacent base pairs stack on opposite sides of the intercalator) [33]. In order to accommodate a small molecule in this mode of binding the DNA duplex must locally unwind the double helix, which results in the lengthening of the duplex structure.

The second common binding mode is known as groove binding, a mode that was first elucidated by Wartell and Wells et. al. in 1974 [34]. As the result of absorption difference spectra and sedimentation analysis, these researchers reported that netropsin binds to supercoiled DNA without unwinding its double helix, which lead to their suggestion of “minor groove binding” as the likely mode of binding. Minor groove binders are generally formed from several aromatic ring systems such as pyrole, imidazole or furan, connected by bonds that have rotational freedom. These molecules generally have curvature in their structure that enables them to fit into the minor groove and to follow the curvature of the groove [35-38]. These molecules interact with nucleic acids through the H-bonding, Van der Waals interactions and hydrophobic forces, and generally they cause very little perturbation in the DNA structure.

The third mode of binding, which was discovered only recently by Komeda and Williams et. al., is called “backbone tracking” [39]. By X-ray crystallography these researchers have shown that the tetra-amine groups of cytotoxic platinum(II) complex bind to DNA by making N-O-N hydrogen bonds with oxygen atoms of the phosphate backbone. Structural deformation of the B-form DNA structure was observed upon ligand binding. Only intercalation of nucleic acids will be discussed further in the next section due its relevance to the context of this thesis.

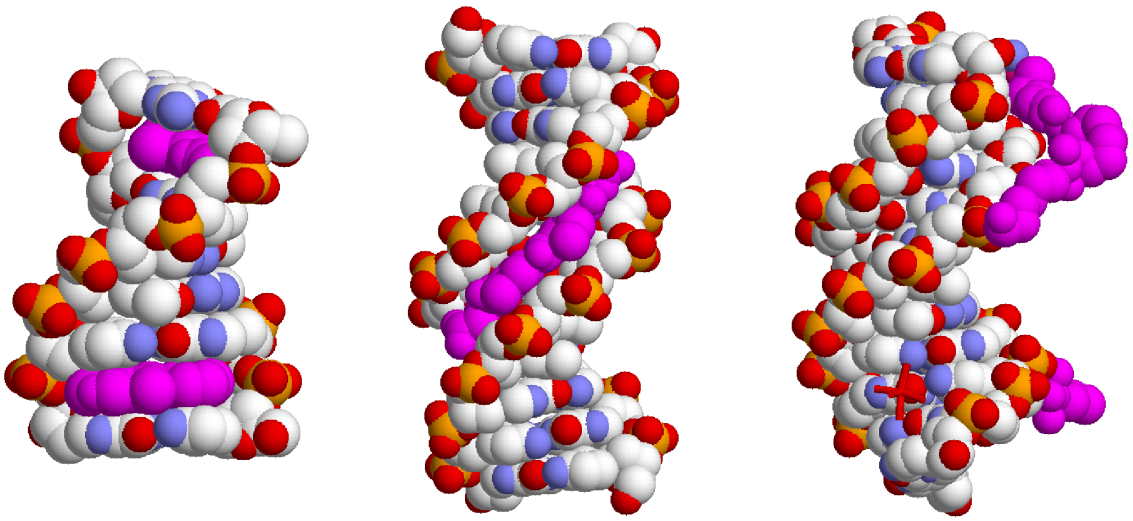

Figure 1.7: The three different binding modes of small molecules (violet) to DNA. Intercalation of ellipticine (left), groove binding of netropsin (middle), and backbone tracking of triplatinNC (right). Picture is created by using PDB ids 1zf3 [40], 121d [41]and 2dyw [39] respectively.

1.3.2 Intercalation

Nucleic acid intercalators are molecules with planar heterocyles that are approximately the same size and the shape as a Watson-Crick base pair. These molecules

generally carry at least one positive charge. The most common classes of intercalators are acridines, anthraquinones, phenanthridinium ions and quinolizinium derivatives (Figure 1.8). The insertion of these heterocyles between the DNA bases of a duplex is driven by electrostatic interactions, dipole-dipole interactions, dispersive interactions and π-π stacking. The binding of an intercalator is generally enthalpically driven as opposed to being entropically driven (as is often the case for groove binders). The association constants of small molecule intercalators for DNA are generally in the range of 10^4-10^6 M^{-1} [36, 42-44].

Intercalation has been established as an anti-cooperative, two stage process. The first and faster of the two stages is the diffusion of the intercalators to the outside of the double helix. This stage is followed by the second, slower stage, which involves the insertion of the intercalator between the base pairs. Intercalation generally takes place on the millisecond time scale [7]. Of course, for the intercalator to be inserted between the base pairs, the DNA needs to unwind during this binding process. The DNA helix lengthens about 3.4 Å, i.e. the thickness of one stacked π system, a structural transition that often results in changes in sugar puckering [7, 44]. In addition, due to the increased distance between the phosphate groups, the charge density of the whole duplex in affected, which is expected to result in cation release from the double helix [7].

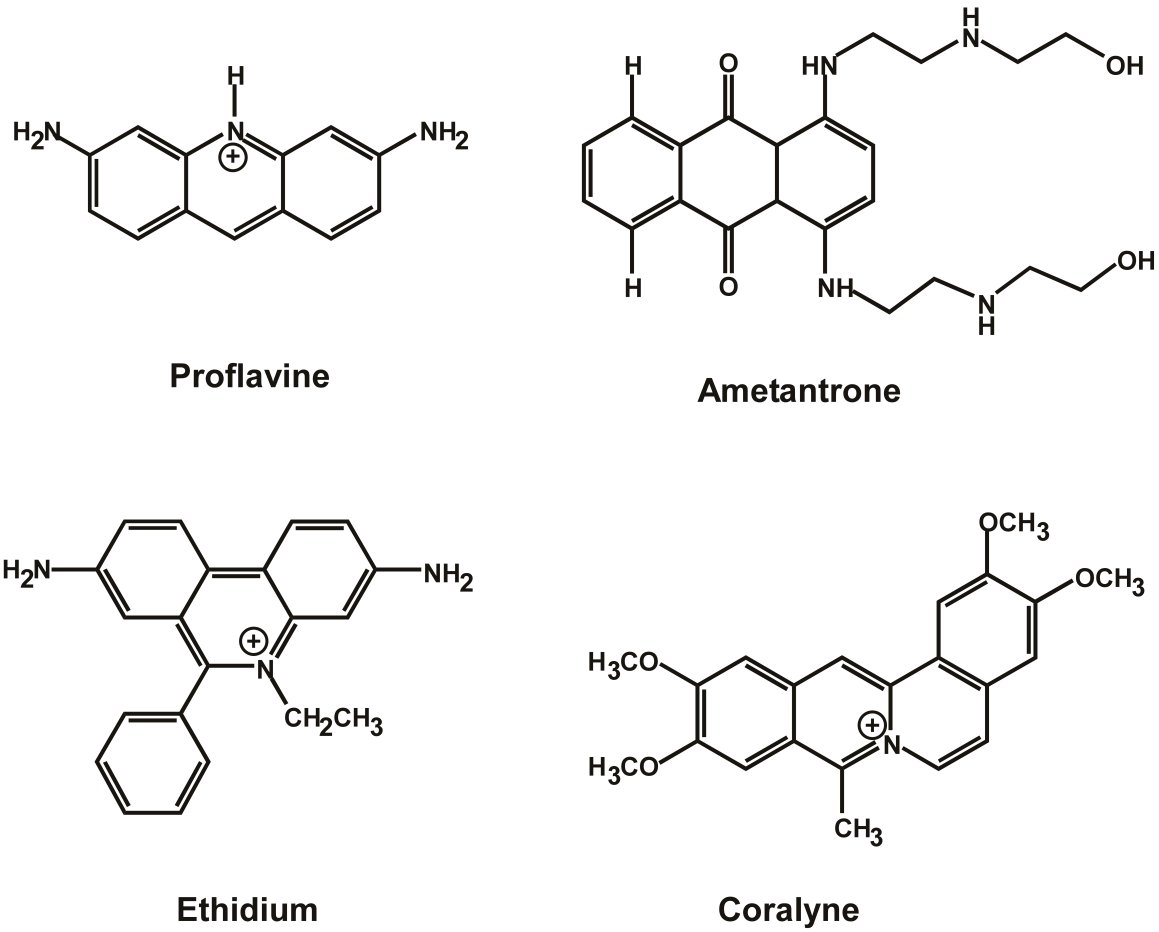

Figure 1.8: Structures of intercalators belonging to different classes: proflavine (acridine), amentantrone (anthraquinones), ethidium (phenanthridinium) and coralyne (quinolizinium).

Intercalators bind to nucleic acids according to "the nearest neighbor exclusion principle" [7, 36]. According to this principle, the binding of an intercalator between the base pairs prevents the binding of another intercalator between the next base pair, or neighboring binding site. In other words, small molecules can intercalate only between every other base pairing in a DNA duplex. Binding isotherms of intercalators that first revealed the nearest neighbor exclusion principle were calculated by Crothers [45]. Structural confirmation of this principle came in 1975 from X-ray fiber diffraction studies

by Lippard et. al. [45]. Today the nearest neighbor exclusion principle is considered as common knowledge by nucleic acid chemists and biophysicists. However, the factors underlying this principle are still unclear [7, 36]. That is, the actual cause of an intercalating molecule not being able to bind to the neighboring binding site to one that is already occupied by intercalator is not known; whether it is steric factors, change in DNA structures upon binding of the first molecule or electrostatic changes in DNA structure, or even repulsion between two closely-spaced intercalators.

1.4. NUCLEIC ACID ASSEMBLY BY INTERCALATION

1.4.1 Molecular Midwife Hypothesis

The "molecular midwife hypothesis" was proposed by Hud and Anet in 2000 as a possible solution to the abiotic formation of the first polymeric nucleic acids, or proto-RNA [46]. The term molecular midwife refers to the any small molecule that is able to facilitate the formation and replication of RNA-like molecules through non-covalent interactions. In the hypothesis put for by Hud and Anet there are two main roles suggested for the molecular midwife; first, to act as a template for the assembly of the nucleic acid bases and/or base pairs into columnar stacks before they are stitched together by a common backbone and, second, to alter the secondary structure of proto-RNA such that replication can take place without the aid of macromolecular enzymes [46].

The assembly of nucleic acid bases by a molecular midwife has been suggested to have two main advantages for proto-RNA replication. First, different bases and/or base pairs might interact most favorably with a particular molecular midwife, resulting in the

selection of certain base pairs from a mixed solution of bases and a multitude of possible pairings. One can imagine different molecular midwives for different base pairs (e.g. Watson-Crick versus Hoogsteen), or even different pairing numbers (e.g. base pairs versus base tetrads) (Figure 1.9). In addition, molecular midwife-mediated assemblies are solely based on reversible non-covalent interactions; particular base pairs could be selected profoundly by the thermodynamics of assembly (i.e the lowest energy structure prevails). Second, the stacking of bases onto a midwife will increase the local concentration of the bases at a certain place, which would certainly promote base (or nucleotide) polymerization under dilute solution conditions [46].

Hud and Anet originally proposed that the structure of the prebiotic molecular midwife was similar to a phthalocyanine, a molecule with a shape and size similar to base tetrads. It was also proposed that the midwife molecule needed to be cationic; otherwise such a molecule would tend to self-associate in solution. In this respect, understanding the interactions of small molecules with nucleic acids, the molecular recognition and assembly of nucleic acid structures by small molecules could represent new opportunities for understanding the formation of the first proto-RNA structures [46], and provide a practical means to assembly and replicate RNA-like polymers for use in materials and biotechnology applications.

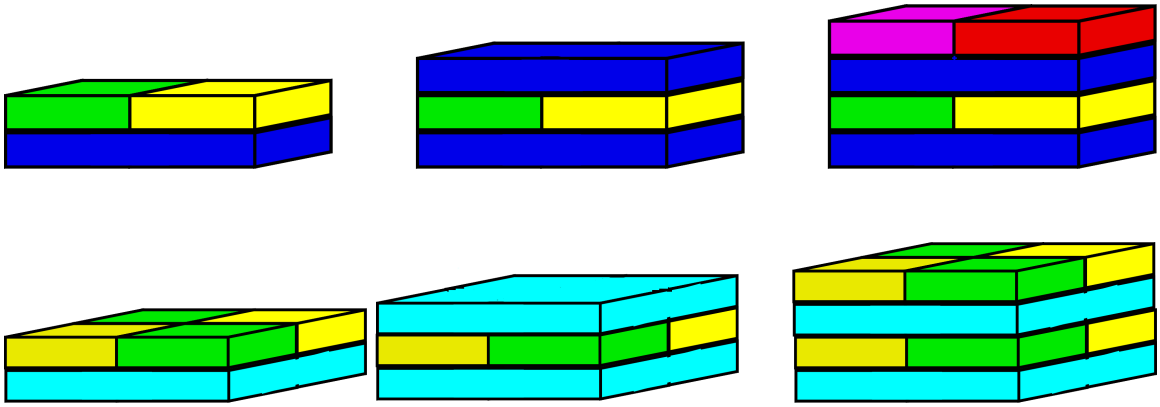

Figure 1.9: Schematic representation of the 'molecular midwife' hypothesis. A midwife molecule (in blue at top, cyan below) acts as a template for the formation and assembly of base pairs. Formation of different base pairing arrangements might be selected based on the structure of the "midwife". In top scheme the blue template is able to assemble base pairs (shown in yellow-green and red-violet). At the bottom, the cyan template, which has a larger surface, is able to assemble base tetrads rather than base pairs. Figure is adapted from [46].

Once the correct bases are selected and assembled into columnar stacks by midwife molecules, the bases of adjacent pairs (or tetrads) on the opposite sides of a midwife will be separated only by 6.8Å, as in an intercalated dinucleotide structure [47]. This arrangement of the bases is ideal for subsequent coupling into polymers by backbone linkages (i.e. phosphodiester linkages). When proto-RNA or RNA-like polymers are formed as part of a midwife-nucleotide stack these polymers will have secondary structures that resemble contemporary RNA and DNA duplexes intercalated by planar molecules (e.g. fluorescent dyes). Because the interactions of small molecules with the proto-RNA structure are non-covalent, one can imagine the removal of intercalators easily from an intercalated proto-RNA structure under appropriate solution conditions. If the proto-RNA secondary structure is only stable when intercalated, one can imagine that the dissociation of the small molecules will leave proto-RNA as single

stranded. This structural transition would represent an important step towards proto-RNA replication (i.e. strand separation). In this respect, understanding the role of intercalators in assembling nucleic acids to different secondary structures which are not stable otherwise, or altering the secondary structures of nucleic acids is an important step towards achieving artificial self-replication of RNA-like polymers.

One example of a molecular midwife has already been demonstrated by the Hud laboratory in a template directed synthesis reaction [48, 49]. The midwife is a molecule known as proflavine that was shown to assemble single strands of 3′-phosphorothioate-$(dT)_3$ and 5′-iodo-$(dT)_4$ onto $(dA)_7$ template and cause the formation of a nicked duplex that is unstable in the absence of the proflavine. The assembled strands were able to ligate resulting in a more stable duplex formation. The selectivity of proflavine as a midwife was also demonstrated in this study. Proflavine, which has a similar structure to Watson-Crick base pairs, was able act as a molecular midwife and ligate two strands that were assembled into a duplex prior to ligation. Another intercalator, coralyne, which is known to preferentially intercalate triplex structures, was not able to promote ligation in the duplex system. These results demonstrated not only that the assembly of nucleic acids can promoted by intercalation, but also the role of intercalator in selecting a particular nucleic acid secondary structure.

1.4.2 Intercalator mediated assembly in Nanotechnology*

Progress in the development of nanotechnology is being advanced by the use of idea inspired by and materials from living organisms. Following the vision of Seeman [50], DNA has emerged as the material of choice for the construction of nanometer-scale objects. Branched DNA structures have proven particularly useful as structural elements in nanotechnology, from the early construction of a 14-faced truncated octahedron [51], to the more recent DNA smiley faces [52]. The relative ease with which a wide variety of molecules can be covalently attached to the termini of oligonucleotides has also enabled the use of DNA in the rational design of new sensors and materials [53]. Although developments in DNA nanotechnology continue to progress at an exciting pace, the use and the role of non-covalent interactions of small molecules, in this case intercalation, in creating such molecular assemblies is relatively unexplored. The intercalation of DNA by small molecules can add another dimension to the DNA base nanotechnology in two different aspects.

The first one is, as discussed in the previous section, the use of small molecules, 'molecular midwifes' to drive the assembly of nucleic acid structures and to alter the secondary structures of nucleic acids. In fact, the first DNA-based nanomechanical device, developed in 1998, was based on the intercalation of ethidium to induce a structural transition in a DNA structure [54].

---

* The work presented in here was previously presented at greater length in Persil et. al. *Trends in Biotechnology*, 2007, 25, 433-436.

The second use of intercalation is in the supramolecular assembly of small molecules into ordered arrays by using DNA as a template. Recently, Armitage et. al. reported that the branched DNA structures can be used to concentrate fluorescent dyes to create supramolecular assemblies which they call “nanotags” [55]. The increased local concentration of the dyes results in extremely bright fluorescent probes. These probes should prove useful in many laboratories, given the ease by which absorbance wavelength, emission wavelength and intensity can be adjusted by tailoring both DNA and dye structures. In general, there is a bright future for DNA based nanotechnology and small molecule binding to nucleic acids will add a new dimension to nanotechnology and bring in applications such as the development of new fluorescent labels or nanomechanical devices.

CHAPTER 2

ASSEMBLY OF AN ANTIPARALLEL HOMO-ADENINE DNA DUPLEX BY SMALL MOLECULE BINDING*

2.1. INTRODUCTION

Molecules that bind DNA and trigger the formation of non-Watson-Crick secondary structures would be useful in the design of dynamic DNA nanostructures [56, 57] and as potential leads for new therapeutic agents [58]. The small molecule coralyne (Figure 2.1) was found to disproportionate polyd(T)·polyd(A) duplex into equal molar equivalents of polyd(T)·polyd(A)·polyd(T) and polyd(A) [59]. By using spectroscopy it has been reported that polyd(A) forms a self structure in the presence of coralyne. Induced band formation is observed upon coralyne binding to polyd(A) in CD spectra (Figure 2.2). Monitoring the change in CD spectra with varying temperature revealed a T_m of 47 ^{o}C. Furthermore, job plot analysis reveal a binding of one coralyne to four adenine bases (Figure 4.2) [59]. Later on same spectral changes were also reported with $d(A)_{32}$ and $d(A)_{16}$ upon binding of coralyne to these oligonucleotides [37].

---

* The majority of the work presented in this chapter was previously published in Persil, Ö., et al., *J. Am. Chem. Soc.*, **2004,** *126*, 8644-8645.

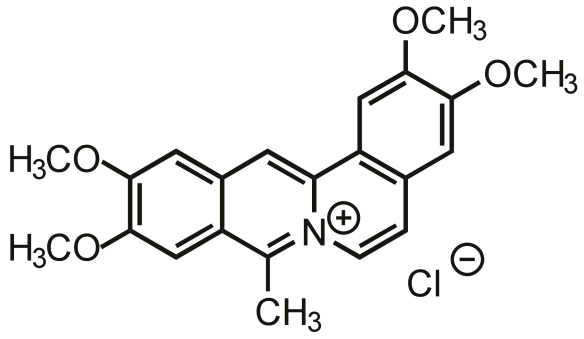

Figure 2.1: Coralyne Chloride Structure

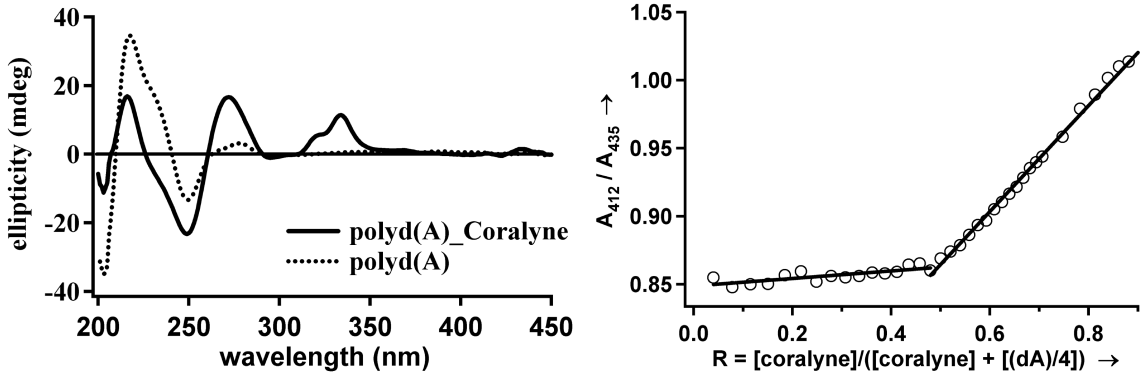

Figure 2.2: CD spectra (left panel) of polyd(A) in the presence (solid line) and absence (dotted line) of coralyne at 6 °C. Samples were 55 µM in adenine base and 14 µM coralyne. Job plot analysis of polyd(A) with coralyne at 22 °C. A_{412}/A_{435} is the ratio of absorbance of coralyne at 412 nm versus 435 nm. The total concentration of coralyne and polyd(A) (in nucleotides / 4) was held constant at 50 µM over the course of titration. Both samples were in 13 mM NaCacodylic buffer with 115 mM NaCl at pH 6.8. Adapted from [59]

These observations suggested that poly(dA) forms a coralyne-dependent secondary structure. However, the basic features of this putative homo-adenine structure, such as strand number (e.g., duplex vs quadruplex) and strand orientation (i.e., parallel vs antiparallel), were not obvious. Here we demonstrate that coralyne promotes the

formation of an antiparallel homo-adenine duplex. Furthermore, the helix of this duplex is compatible with flanking Watson-Crick helices. To the best of our knowledge, this represents the first report of a homoadenine duplex at neutral pH.

2.2. EXPERIMENTAL PROCEDURES

2.2.1 Materials

$d(A)_4$ and 3'-$d(A)_8$ -5'-5'-$d(A)_8$ -3' oligodeoxynucleotides were synthesized in house on an automated synthesizer using standard phosphoramidite chemistry. $d(A)_8$, $d(A)_{16}$, $d(A)_{32}$, d(GACCCGCA_8CCTCGCC) and d(GGCGAGGA_8GCGGGTC) oligonucleotides were purchased from IDT (Coralville, IA). Full-length oligonucleotides were separated from failure sequences by denaturing polyacrylamide gel electrophoresis. Full-length products were extracted from the gel matrix using the crush-and-soak method followed by ethanol precipitation and desalting by passage over a 1 m G-25 sephadex column. Column fractions containing purified oligonucleotides were pooled, lyophilized and resuspended in dH_2O. Coralyne chloride and proflavine hemisulfate were purchased from Sigma and used without further purification.

Oligonucleotide and small molecule concentrations were determined by UV-Vis spectroscopy using the following extinction coefficients: $d(A)_4$, ε_{260} = 51 400 M^{-1} cm^{-1}; $d(A)_8$, ε_{260} = 99 400 M^{-1} cm^{-1}; $d(A)_{16}$, ε_{260} = 195 400 M^{-1} cm^{-1}; $d(A)_{32}$, ε_{260} = 387 400 M^{-1} cm^{-1}; d(GACCCGCA_8CCTCGCC), ε_{260} = 212 100 M^{-1} cm^{-1}; and d(GGCGAGGA_8GCGGGTC), ε_{260} = 231 900 M^{-1} cm^{-1}; 3'-$d(A)_8$-5'-5'-$d(A)_8$-3', ε_{260} = 195 400 M^{-1} cm^{-1}; coralyne chloride, ε_{420} = 14 500 M^{-1} cm^{-1}; proflavine hemisulfate, ε_{444} = 38 900 M^{-1} cm^{-1}

2.2.2 Polyacrylamide Gel Electrophoresis (PAGE)

Oligonucleotides were 5'-end labeled using γ-^{32}P-ATP (MP Biomedical, formerly ICN) and T4 polynucleotide kinase (New England Biolabs). Non-denaturing PAGE experiments were run in a standard 1×TBE buffer (Tris-Borate-EDTA), at a constant power of 7 W and an ambient temperature of 4 °C. Gels were imaged using a Fuji Phosphor Imager (FLA-3000).

2.2.3 Circular Dichroism (CD) and UV-Vis spectrophotometry

CD spectra were acquired on a JASCO J-810 CD spectropolarimeter equipped with Peltier temperature control unit. Spectra were acquired using a 1 cm path length cell. UV-Vis absorbance measurements were performed using a HP 8453 UV-Vis diode array spectrophotometer equipped with an Agilent 89090A Peltier temperature control unit. UV melting profiles were acquired by increasing the sample temperature at a rate of 1 °C min^{-1} from 5 to 80 °C.

2.2.4 Atomic Force Microscopy (AFM)

Scanning force images were acquired using a Nanoscope IIIa AFM (Digital Instruments) equipped with a J scanner operating in tapping mode. AFM tips were NSC12 non-contact silicon rectangular cantilevers (Mikromasch USA, Portland), which were cleaned with ozone prior to use. Samples (20 μl) were deposited onto freshly cleaved mica, incubated for 30 min at 4 °C, rinsed once with dH_2O (4 °C), wicked dry by touching an edge of the mica to filter paper, blown dry with nitrogen gas and stored overnight in a vacuum desiccator at room temperature. Samples were imaged under ambient conditions. Images were digitally flattened to remove background slope in the horizontal dimensions

2.3. RESULTS AND DISCUSSION

2.3.1. Binding of Coralyne to Homo-Adenine Oligonucleotide, $d(A)_n$

We first investigated the binding of coralyne to different lengths of $d(A)_n$ oligonucleotides with CD and UV-Vis spectroscopy. CD spectra shown in Figure 2.3 demonstrate that $d(A)_{16}$ undergoes a change in secondary structure very similar to polyd(A) (Figure 2.3) in the presence of coralyne. The positive CD bands observed between 310 and 350 nm indicate the binding of coralyne within the chiral environment of DNA. These CD bands are quite pronounced for coralyne in the presence of $d(A)_{16}$ and $d(A)_{32}$ at 5 °C (Figure 2.4). However, a sample of $d(A)_8$ containing the same concentration of coralyne and DNA (in nucleotide base) exhibits a lower degree of coralyne binding, and coralyne CD bands are not at all detected with $d(A)_4$ (Figure 2.4).

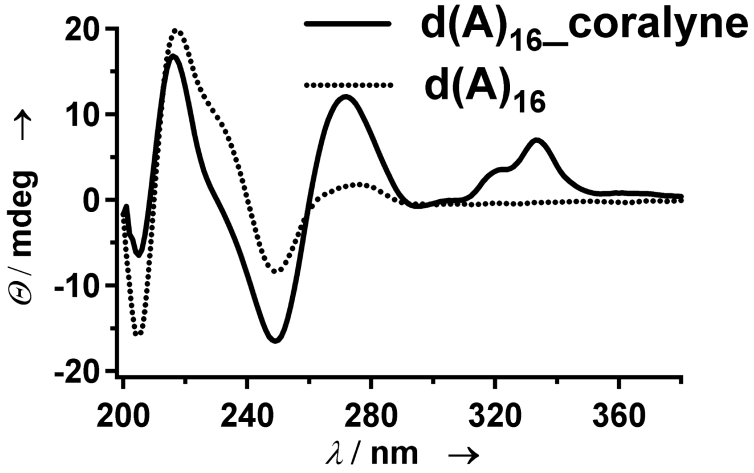

Figure 2.3: CD spectra of $d(A)_{16}$ in the presence (solid line) and absence (dotted line) of coralyne at 5 °C. Samples were 55 µM in adenine base and 14 µM coralyne and in 13 mM NaCacodylic buffer with 115 mM NaCl at pH 6.8.

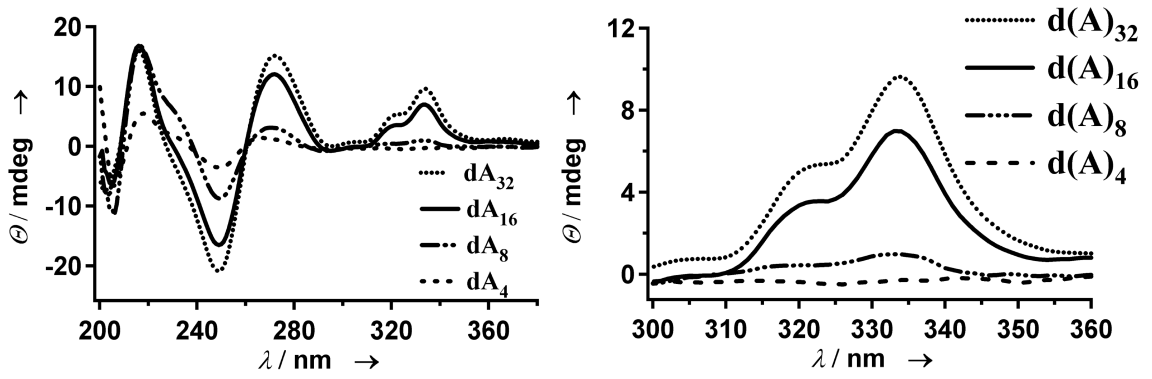

Figure 2.4: CD spectra of $d(A)_n$ oligonucleotides in the presence of coralyne at 5 ^oC. The right panel shows the longer wavelength region of the whole CD spectra (left panel) Samples were 55 µM in adenine base and 14 µM coralyne in 13 mM NaCacodylic buffer with 115 mM NaCl at pH 6.8.

The UV absorption spectra of coralyne was hypochromic and red shifted around 15 nm upon binding to $d(A)_n$ olignonucleotides as represented in Figure 2.5. A hypochromicity and red shift upon small molecule binding to nucleic acids is generally associated with the π-π stacking of the small molecule with the nucleic acid bases, and is therefore considered an indication of the intercalative mode of binding [36]. Monitoring the changes in absorption and CD spectra revealed a T_m of 41 ^oC, 24 ^oC and 11 ^oC for $d(A)_{32}$-coralyne, $d(A)_{16}$-coralyne and $d(A)_8$-coralyne complexes respectively.

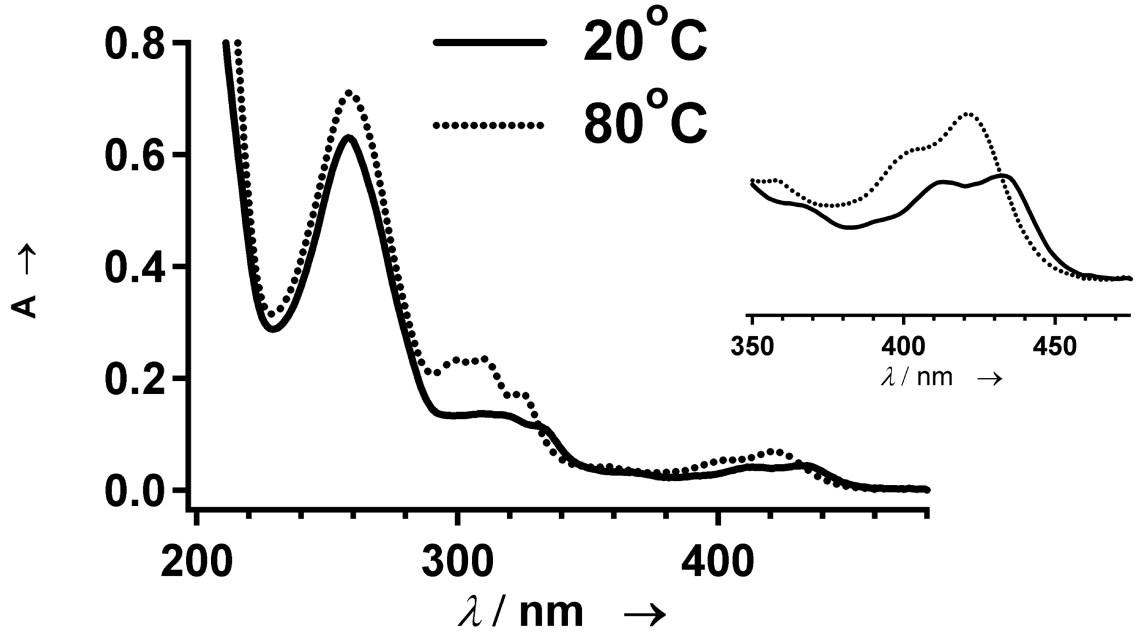

Figure 2.5: UV spectra of $d(A)_{16}$ in the presence of coralyne at 20 °C and 80 °C. The insert shows the longer wavelength region of the UV spectra. Samples were 55 µM in adenine base and 14 µM coralyne in 13 mM NaCacodylic buffer with 115 mM NaCl at pH 6.8.

2.3.2. Effect of pH on Coralyne binding to $d(A)_n$

Poly(rA) is known to form a duplex with parallel strands under acidic conditions [60]. The structure is held by two H-bonds between the adenines and stabilized further by H-bonds between the protonated N1 and the oxygen of the phosphate backbone. In the present study, the effect of pH on coralyne $d(A)_n$ assemblies was investigated by CD spectroscopy. As demonstrated in Figure 2.6, the intensity of the observed induced band is decreased with decreasing pH. In addition, below pH 5 the DNA secondary structure is altered, as apparent from the change observed in 200-300 nm region. We rationalize that this change is due to the protonation of adenine N1, because the pKa of this nitrogen is

around 4.0 [61]. This result suggests the important role of N1 and its possible involvement in an anticipated base pairing.

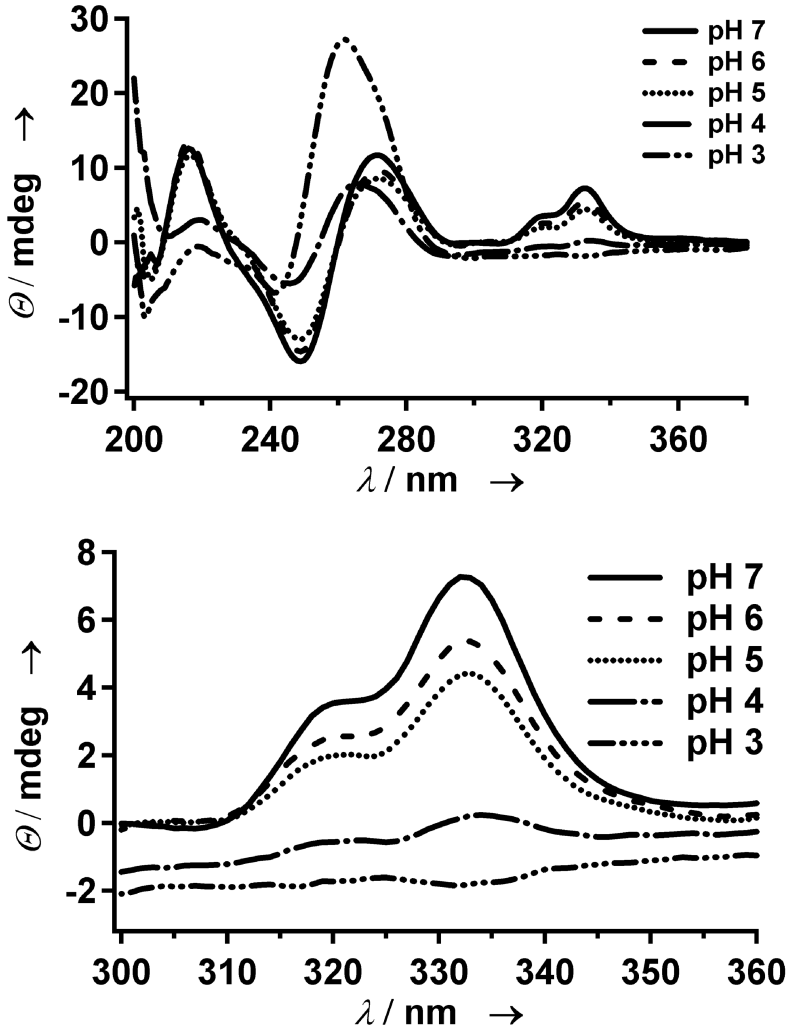

Figure 2.6: The whole (top) and the selected region (bottom) of the CD spectra of $d(A)_{32}$ at varying pH values. Samples with pH values lower than 7 were prepared by the titration of a pH 7 sample with 1 M HCl. All samples were 55 μM nucleotide base, 14 μM coralyne, 115 mM NaCl and 13 mM NaCacodylic buffer pH 6.8.

UV absorption studies demonstrate that the stability of the $d(A)_{16}$-coralyne structure depends on both DNA and coralyne concentration. Figure 2.8 is showing the

plot of percentage of $d(A)_{16}$-coralyne complex formed as a function of DNA and coralyne concentration. In this study, a sample of poly$d(A)_{16}$ 200 µM in nucleotide base and 50 µM in coralyne was diluted with buffer so that the ratio of $d(A)_{16}$ to coralyne was kept constant over the course of the dilution. UV absorption spectra were collected after each dilution. The percentage of $d(A)_{16}$-coralyne complex formed at each point was determined by performing a least-squares fit of the corresponding UV absorption spectrum as a weighted sum of two absorption spectra, which were the spectrum of 200 µM $d(A)_{16}$, 50 µM coralyne (bound coralyne) and 0.025µM coralyne (free coralyne) (Figure 2.7).

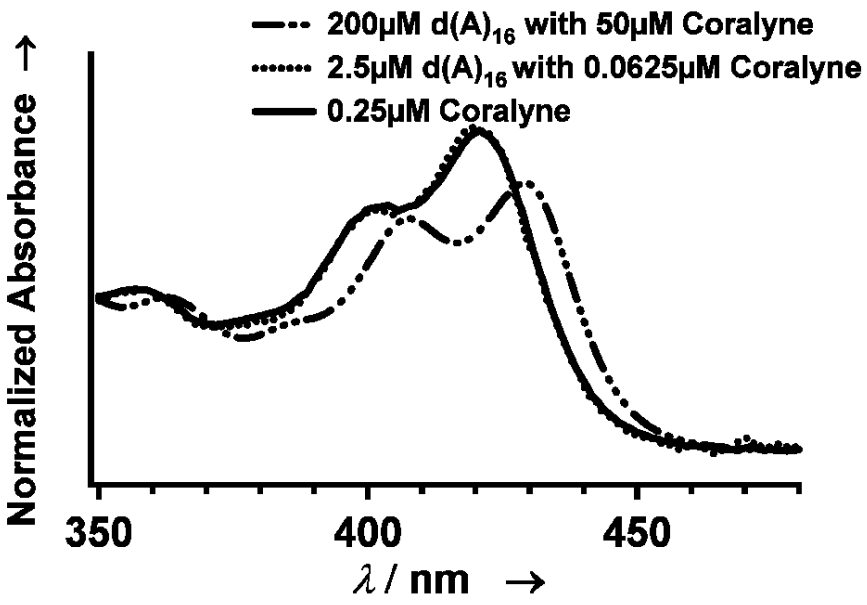

Figure 2.7: 340-480 nm region of UV absorbance spectra of coralyne in the absence and presence of $d(A)_{16}$ at room temperature. These are representative of UV absorbance spectra used to generate the plot in Figure 2.8. Samples were in 115 mM NaCl and 13 mM NaCacodylic buffer at pH 6.8.

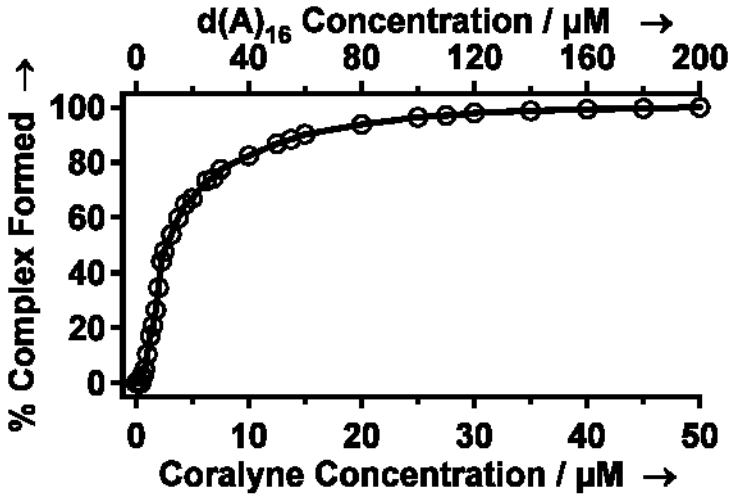

Figure 2.8: Plot of percentage of $d(A)_{16}$-coralyne complex as function of DNA and coralyne concentration. A sample of poly(A)-coralyne complex which is 200 µM in adenine base and 50 µM in coralyne in 13 mM NaCacodylic and 115 mM NaCl is diluted with 13 mM NaCacodylic buffer and 115 mM NaCl. A spectrum is taken after each dilution at room temperature, and the spectra fitted to a two state model.

2.3.3. Morphology and the Strand Polarity of $d(A)_n$-Coralyne Assemblies

We investigated the binding of coralyne to different lengths of homo adenine oligonucleotides, $d(A)_n$ also with non-denaturing polyacrylamide gels (PAGE). However, in order to be able to observe the effect of coralyne on $d(A)_n$ oligonuclotides, we needed to cast the coralyne also inside the gel. Our initial attempts running $d(A)_n$-coralyne samples directly on PAGE didn't reveal any change on mobility, where all the $d(A)_n$ oligonucleotides run as single stranded DNA. The reason is most likely the mobility of coralyne (positively charged) and nucleic acid (negatively charged) in different directions towards different electrodes as soon as the electric field is applied. Therefore, by casting the coralyne inside the gel, we provide an environment where the $d(A)_n$ strands will always be in equilibrium with coralyne. In Figure 2.9 we present an image of a nondenaturing polyacrylamide gel cast with 20 µM coralyne in which the

oligonucleotides $d(A)_4$, $d(A)_8$, $d(A)_{16}$, and $d(A)_{32}$ were subjected to electrophoresis. Of these four oligonucleotides only $d(A)_4$, the shortest of the series, migrates with the mobility expected for a single-stranded oligonucleotide. Oligonucleotides $d(A)_8$, $d(A)_{16}$, and $d(A)_{32}$ apparently form complexes with coralyne that are too large to migrate appreciably through the gel matrix.

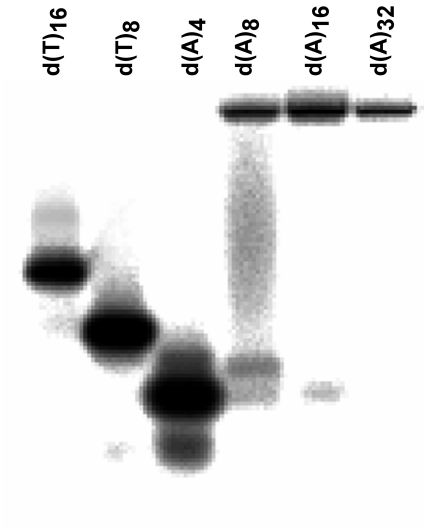

Figure 2.9: Gel mobility experiment for $d(A)_n$ oligonucleotides in a nondenaturing 15% polyacrylamide gel cast with 20 µM coralyne (run at 4 °C ambient temperature). $d(T)_8$ and $d(T)_{16}$, which do not bind coralyne, serve as molecular weight markers. Samples loaded into the gel were 16 µM in nucleotide base and 16 µM in coralyne

To directly examine the size and morphology of these large $d(A)_n$-coralyne complexes, $d(A)_n$ oligonucleotides in solution with coralyne were deposited on mica and imaged using an atomic force microscope (AFM). A representative image from the $d(A)_{16}$ preparation is shown in Figure 2.10, which reveals polymers of micrometers in length but only 1.6 nm (±0.3 nm) in height. Similar polymers were observed for $d(A)_8$ and $d(A)_{32}$ in the presence of coralyne, but not for $d(A)_4$, or any of the $d(A)_n$

oligonucleotides in the absence of coralyne. The length of these molecular assemblies suggests the staggered pairing of $d(A)_n$ oligonucleotides into quasiinfinite multistranded polymers that minimize exposed ends of unstacked bases and maximize binding sites for coralyne. In addition the height measurements suggest that the assemblies were most likely double helical since the width of the B-form DNA is around 2 nm. And upon intercalation it has been know that the helix unwinds and get elongated [36].

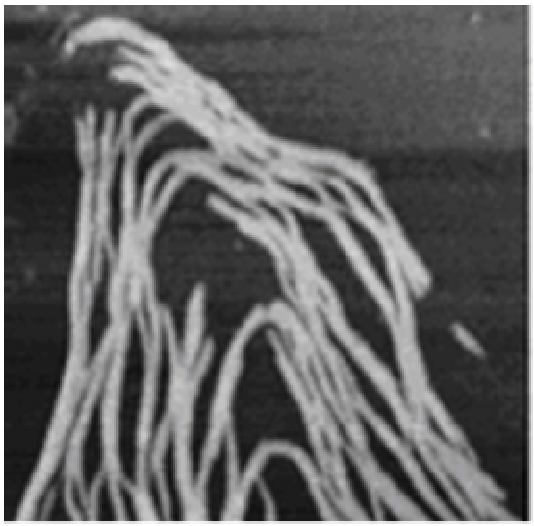

Figure 2.10: AFM image of homo-adenine-coralyne assemblies formed by $d(A)_{16}$ in the presence of coralyne. Samples were 220 µM in nucleotide base, 55 µM coralyne and imaged on freshly cleaved mica after drying. Picture is 800 nm x 800 nm.

To investigate strand polarity in the $d(A)_n$-coralyne structures (i.e. anti-parallel versus parallel) we prepared the oligonucleotide 3'-$d(A)_8$-5'-5'-$d(A)_8$-3', an analogue of $(dA)_{16}$ that differs by a change in strand polarity at the center of the oligonucleotide. If the homo-adenine-coralyne secondary structure contains only parallel DNA strands, then this particular oligonucleotide would be restricted to the assembly of structures only 16 nucleotides in length (Figure 2.11). On the other hand, if the homo-adenine-coralyne

structure contains antiparallel strands, then this oligonucleotide would assemble into quasi-infinite polymers (Figure 2.11). The observation of micrometer-length fibers formed from 3'-d$(A)_8$-5'-5'-d$(A)_8$-3' and coralyne indicates the inclusion of antiparallel strands in this structure (Figure 2.12). A similar antiparallel strand arrangement is also likely to exist in the $(dA)_n$-coralyne assemblies; however, a parallel stranded arrangement could not strictly be ruled out by these experiments.

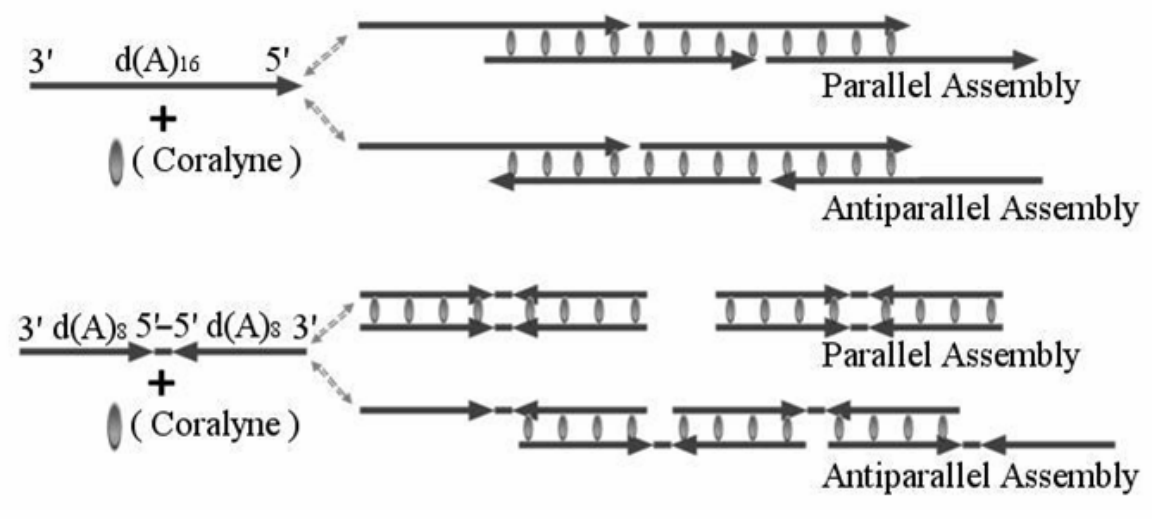

Figure 2.11: Possible assembly constraints implicit to d$(A)_{16}$ and 3'-d$(A)_8$-5'-5'-d$(A)_8$-3' in the presence of coralyne.

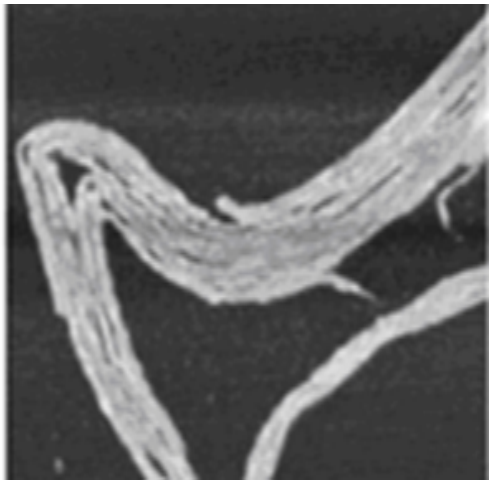

Figure 2.12: AFM image of homo-adenine-coralyne assemblies formed by 3'-d$(A)_8$-5'-5'-d$(A)_8$-3' in the presence of coralyne. Samples were 220 µM in nucleotide base, 55 µM coralyne and imaged on freshly cleaved mica after drying. Picture is 800 nm x 800 nm.

2.3.4. Watson-Crick Compatibility of d$(A)_n$-Coralyne Assemblies

To determine if the $(dA)_n$-coralyne structure is composed of two strands, as opposed to three or more, we tested the ability for the homo-adenine-coralyne structure to be accommodated within a Watson-Crick duplex. This was accomplished by repeating the series of CD and UV-Vis spectroscopy experiments described above with a duplex formed by the oligonucleotides 5'-d(GACCCGC-A_8-CCTCGCC)-3' (**plusA8**) and 5'-d(GGCGAGG-A_8-GCGGGTC)-3' (**minusA8**). These sequences were designed to form two 7-bp Watson-Crick duplexes that are separated by eight A·A mismatches.

CD spectra demonstrate that coralyne binds to the **plusA8·minusA8** duplex in a local environment that is the same as, or very similar to, that of the $(dA)_n$ samples (Figure 2.13). Job plot analysis based upon coralyne absorption measurements established that

coralyne binds to the duplex **plusA8·minusA8** at the level of one coralyne per four adenine bases (Figure 2.13), which is suggestive of an intercalative mode of binding within eight A·A base pairs at the level allowed by the nearest-neighbor exclusion principle.

By competition dialysis, it was previously demonstrated that coralyne preferentially binds to triplex DNA and poly(dA) over duplex DNA, possibly due to its larger surface area [62]. In addition, as mentioned before, by using spectroscopy is has been demonstrated that d(T)·d(A) duplexes disproportionate upon binding of coralyne and formed d(T)·d(A)·d(T)–coralyne and d(A)-coralyne complexes [37, 59]. This result is rationalized as the formation of a more thermodynamically stable d(T)·d(A)·d(T)-coralyne and d(A)-coralyne complexes compared to d(T)·d(A)-coralyne complex. In this respect, it is reasonable for coralyne to bind preferentially to incorporated A·A region rather than the flanking Watson-Crick duplex regions.

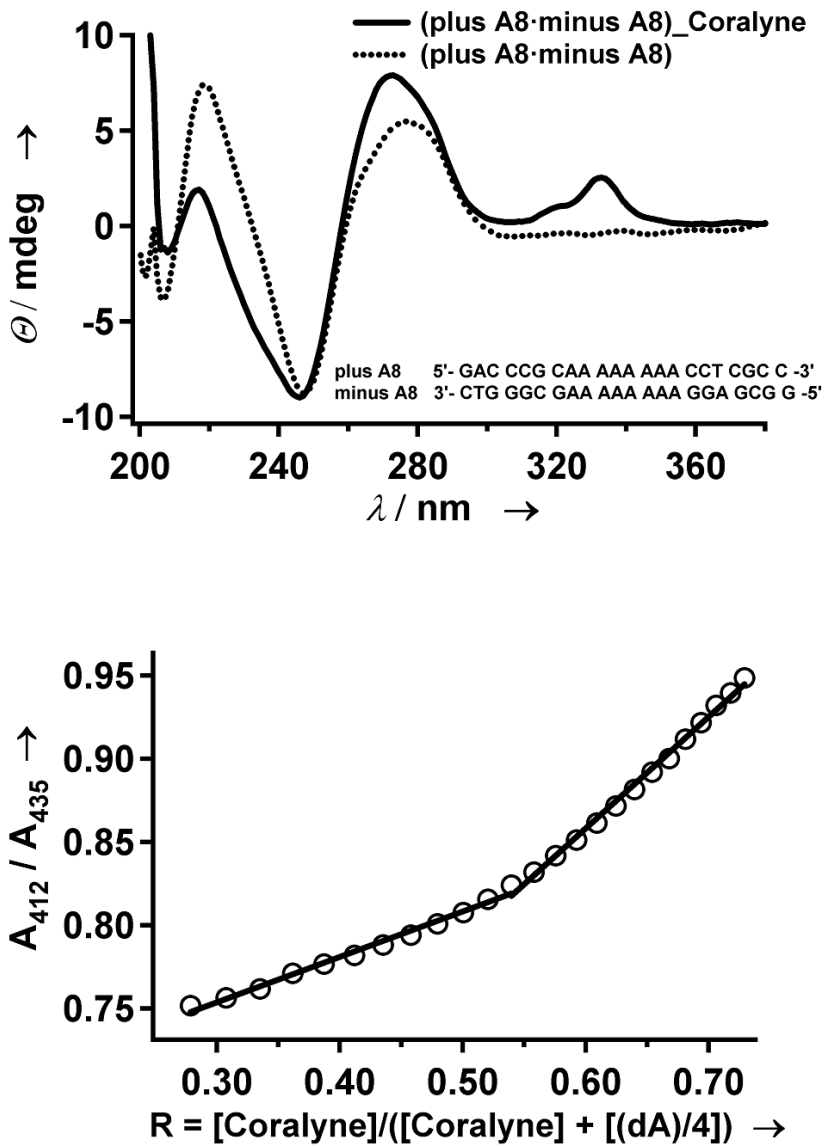

Figure 2.13: CD spectra of duplex **plusA8·minusA8** in the presence and absence of coralyne at 5 ^{o}C (upper panel). Sample was 55 µM in adenine base (only the ones in the middle count) and 14 µM in coralyne. Job Plot analysis of plusA8·minusA8 with coralyne (bottom panel). The combined concentration of coralyne and (dA)/4 was 15 µM for each data point in the Job plot. Both samples are in 13 mM NaCacodylic buffer with 115 mM NaCl at pH 6.8.

Representative UV-Vis absorbance curves for **plusA8·minusA8** duplex in the presence of coralyne at 5 ^{o}C and 80 ^{o}C are given in Figure 2.14. A red shift and

hypochromic effect is observed upon binding of coralyne to **plusA8·minusA8** below the T_m are observed for the $d(A)_n$-coralyne assemblies . UV absorbance curves reveal that the T_m of duplex **plusA8·minusA8** increases by 13 °C, from 44 to 57 °C, upon coralyne binding (Figure 2.15). These results clearly demonstrate that the homoadenine-coralyne structure is compatible with the helical structure of a Watson-Crick duplex. As a comparison, the addition of proflavine, a well-characterized intercalator of Watson-Crick duplexes, only stabilizes the **plusA8·minusA8** duplex by 2 °C (T_m = 46 °C) (Figure 2.15).

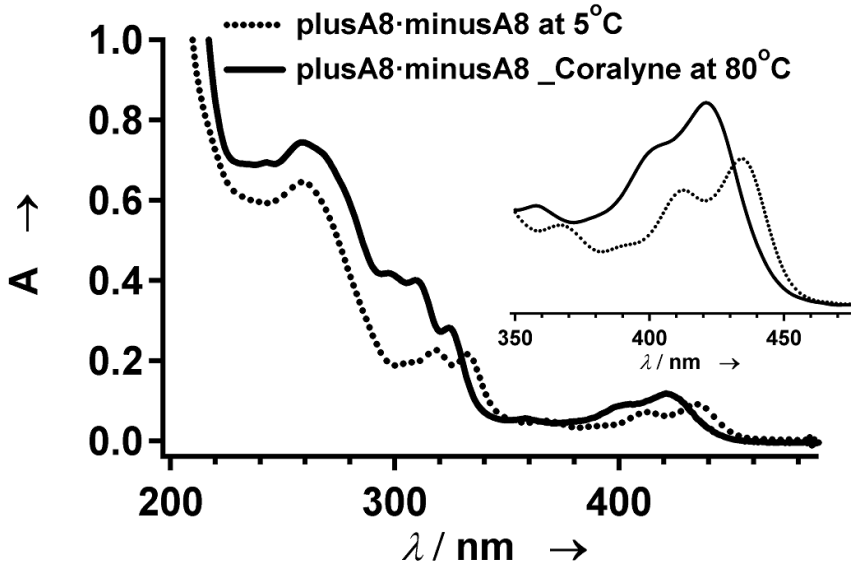

Figure 2.14: UV spectra of duplex **plusA8·minusA8** in the presence of coralyne at 5 °C and 80 °C. The inset shows the coralyne region from 350-480 nm. Sample was 55 µM in adenine base (only the ones in the middle count) and 14 µM in coralyne in 13 mM NaCacodylic buffer with 115 mM NaCl at pH 6.8.

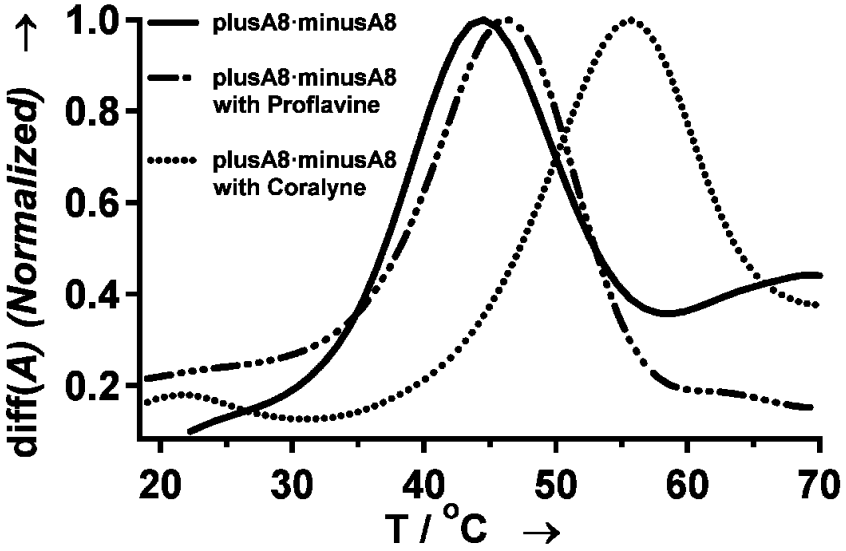

Figure 2.15: Derivative UV melting curves for duplex **plusA8·minusA8** in the absence and presence of coralyne and proflavine, respectively. Samples were 55 µM in adenine base (only the ones in the middle count) and 14 µM in small molecule in 13 mM NaCacodylic buffer with 115 mM NaCl at pH 6.8.

2.4. CONCLUDING REMARKS

In conclusion, we have presented evidence that an antiparallel duplex is formed between homo-adenine sequences in the presence of coralyne. We have also demonstrated that this homoadenine-coralyne structure can be incorporated into duplex structures that include Watson-Crick pairs. Poly(rA) is known to form a duplex with parallel strands and protonated adenine bases under acidic conditions [60]. However, the homo-adenine-coralyne duplex is likely to be significantly different from the low-pH poly(rA) duplex, because the strand orientation has been shown here to be antiparallel and, as opposed to poly(rA), stability decreases as pH is lowered from neutrality. Efforts

are currently underway in our laboratory to determine the exact structure of the A·A base pairing that exists within the homo-adenine-coralyne duplex.

The results presented here are complementary to previous reports that small-molecule binding can be used to alter nucleic acid secondary structure [29, 31, 37, 59], yet go still further to show that small-molecule binding can even be used to drive the formation of non-Watson-Crick duplexes. Furthermore, coralyne in here is an example for the selective molecular recognition of nucleic acids by small molecules and represents the selectivity of the “midwife” for certain base pairing. Coralyne, but not proflavine, is able select plausible non-canonical A·A base pairing and promote the assembly.

CHAPTER 3

PLAUSIBLE A·A STRUCTURE

3.1. INTRODUCTION

DNA secondary structures are stabilized mainly by hydrogen bonds and stacking interactions between base pairs [4]. And the polymorphic structure of DNA give rise to conformational diversity in DNA secondary structures which play important role in the biological processes [29]. Recently, interest especially in non canonical nucleic acid structures has arisen such as triplex or quadruplex DNA as potential therapeutic targets [28]. Using the molecular recognition by small molecules provides unique opportunities in affecting these diverse conformations of DNA [30]. Unique examples of such molecular recognition is reported for conversation of Z-form DNA to intercalated right handed form by the intercalation of ethidium bromide and daunomycin under solution conditions that otherwise favor Z-form DNA structure [63, 64]. Later Chaires et. al. reported not only the allostery of the DNA but the structure of the small molecule is important in such molecular recognition [31]. They have shown that while daunorubicin binds preferentially to right handed DNA, its (-) enantiomer WP900 preferentially binds to left handed DNA. Therefore, there is interest in the design and synthesis of such molecules binding to particular DNA structures selectively and tightly as well as affecting the conformation of these particular DNA sequences in order to affect their biological function [28, 29].

In chapter 2, such a unique selective molecular recognition of homod(A) sequences by coralyne, a small, crescent shaped molecule has been described [65]. We

have reported that coralyne not only binds to homod(A) oligonucleotides selectively and tightly but also promotes the formation of an anti-parallel homod(A) duplex at neutral pH. The formation of this duplex is highly dependent upon the pH and the length of the homo-adenine oligonucleotides. Furthermore, we have illustrated that the helical structure of the oligo-dA-coralyne structure is compatible with the B-form DNA helix [65].

We have reported that this homo-adenine-coralyne duplex is also significantly different from the low-pH poly(rA) duplex. The structure of poly(rA) under acidic conditions, below pH 4.5, has been determined by x-ray fiber diffraction by Rich et. al. in 1961 [60]. In this structure, two adenine strands are organized parallel to each other around a two fold rotation axis. The two adenine strands are mainly held by two hydrogen bonds (H-bonds) between the N6-amino groups and N7 of adenines at opposite strands. The structure has been shown additionally to be stabilized by both the protonated $N1^{+}$-H and N6-amino making a hydrogen bond to the negatively charged phosphate group of the opposite ribose backbone. Sites for the protonation of adenosine are given in Figure 3.1. In here we investigated the duplex formed by the anti-parallel assembly of adenines upon coralyne binding by chemical probing to identify this unique putative A·A base pairing between two adenines at neutral pH.

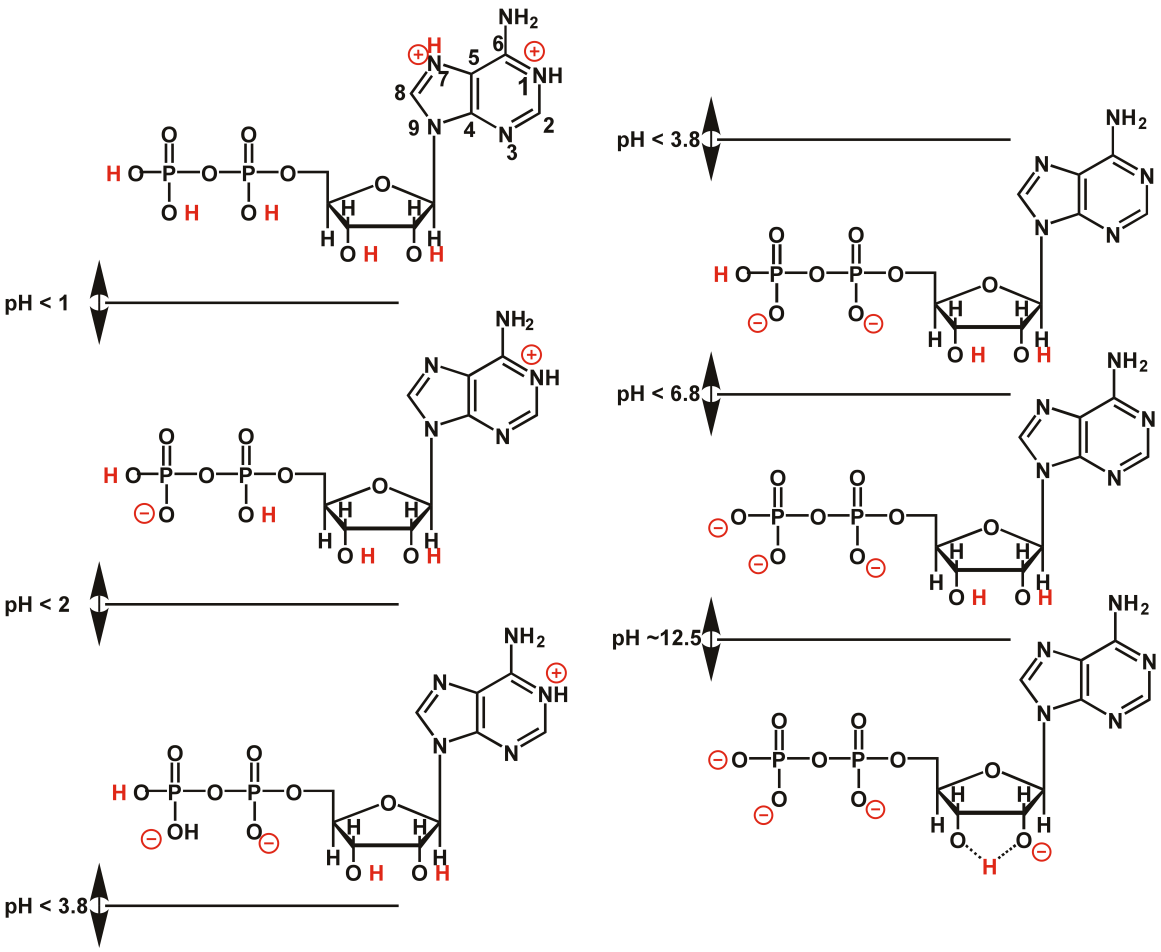

Figure 3.1: Protonation sites in adenosine diphosphate (ADP). Numbering of the adenine base is given in the first ADP structure. Picture is adapted from [7].

3.2. EXPERIMENTAL PROCEDURES

3.2.1 Materials

Coralyne Chloride is purchased from Sigma and the concentrations of stock solutions in H_2O are determined spectrophotometrically by using $\varepsilon_{420} = 14500\ M^{-1}\ cm^{-1}$. The unmodified oligonucleotides were purchased from Integrated DNA Technologies (IDT). Modified oligonueclotides were synthesized on an Expedite 8909 DNA synthesizer. 7-Deaza-dA-CE Phosphoramidite is purchased from Glen Research and

used as received. All of the oligonucleotides were purified by HPLC, desalted and characterized by mass spectroscopy. Concentrations of the stock solutions were calculated by the using the following extinction coefficients supplied by IDT for the unmodified oligonucleotides: **(A)** 5'-d**(**GAC CCG CAA AAA AAA CCT CGC C)-3' ε_{260} = 212 100 M^{-1} cm^{-1} ; **(B)** 5'-d(GGC GAG GAA AAA AAA GCG GGT C)-3', ε260 = 231 900 M^{-1} cm^{-1}. All the samples are 60 µM in adenine base (only the adenines in the middle counted) and 15 µM coralyne. The samples were prepared in 20 mM NaCl and 1X BPE (1 mM Na_2EDTA, 6 mM Na_2HPO_4 and 2 mM NaH_2PO_4) at pH 7.

Oligonucleotide sequences used; 7 refers to 7-deazadeoxyadenosine nucleotide;

(A) 5'-d(GAC CCG CAA AAA AAA CCT CGC C)-3'

(1A) 5'-d(GAC CCG CAA AA7 AAA CCT CGC C)-3

(2A) 5'-d(GAC CCG CAA 7A7 AAA CCT CGC C)-3

(3A) 5'-d(GAC CCG CAA 777 7AA CCT CGC C)-3

(B) 5'-d(GGC GAG GAA AAA AAA GCG GGT C)-3'

(1B) 5'-d(GGC GAG GAA A7A AAA GCG GGT C)-3'

(2B) 5'-d(GGC GAG GAA A7A 7AA GCG GGT C)-3'

(3B) 5'-d(GGC GAG GAA 7A7 AAA GCG GGT C)-3'

4(B) 5'-d(GGC GAG GAA A77 AAA GCG GGT C)-3'

3.2.2 Circular dichroism (CD)

CD spectra were acquired on a JASCO J-810 CD spectropolarimeter equipped with a Peltier temperature control unit. Spectra were acquired using a 5 mm strain-free rectangular cell. CD melting profiles were acquired as full spectra from 5 to 95 ^{o}C in 1 ^{o}C

steps, at 1 ^{o}C/minute. Melting curves were obtained by fitting each spectrum to a two-state model of a linear superposition of the 5 ^{o}C and 95 ^{o}C spectra.

3.3. RESULTS AND DISCUSSION

3.3.1 Rationale behind Chemical Probing

In order to understand the putative A·A base pairing, first we considered all possible base pairing arrangements that can form between two adenine nucleotides at neutral pH theoretically. We disregarded the protonated structures formed by the protonation of N1 under acidic conditions, pH < 4.5 [7] since we know that homodA-coralyne structure do not exist below pH 5. We end up with seven possible non-protonated base pairing arrangements which all happened to be observed in ribosome structures (Figure 3.2) [13, 66-68]. In all of these structures, the involvement of the N6-amino group is a necessity since it is the only H-bond donor in adenine nucleotide.

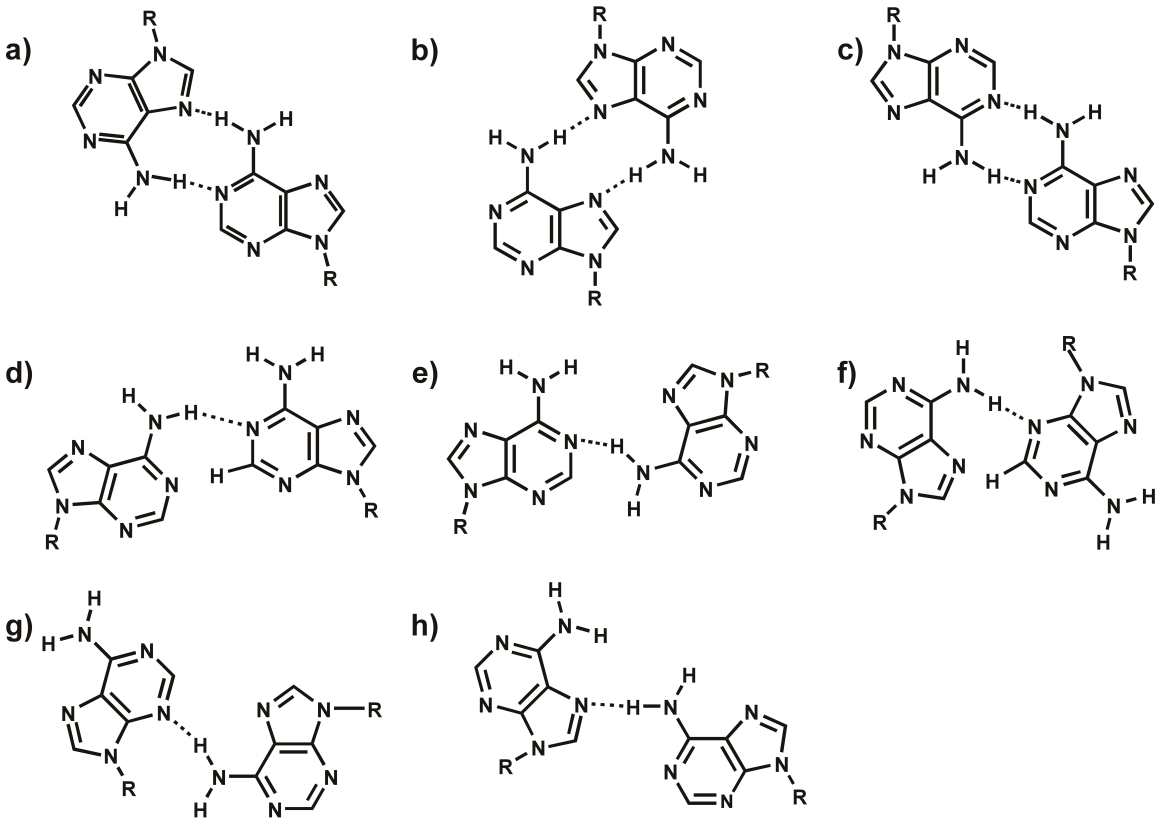

Figure 3.2: Non-protonated possible A·A base pairing arrangements.

Among those possible base pairings only three of them had two H-bonds between the two adenine bases (Figure 3.2). In base pairing 1a, both N1 and N7 located at opposite strands make H-bond to N6-amino adenine located on the opposite strand. In 1b and 1c, only either N7 or N1 is involved in a symmetric H-bonding with N6-amino on the opposite strand. Consequently, in these base pairing arrangements one of the adenine bases need to adopt syn conformation in order to be able to form an anti-parallel duplex. The possible one H-bonded structures are also given in Figure 3.2. Even though we took one H-bonded structures into considerations, we reasoned the putative A·A base pairing had two H-bonds due to achievable greater stability. In addition the frequency of occurrence of the single H-bonded A·A base pairs in ribosome structures is negligible

compared to two H-bonded structures. Accordingly, since we knew that the putative A· A base pairing in our system is disturbed by the protonation of N1, we focused our studies to the involvement of N7 in this A· A base pairing.

3.3.2 Effect of 7-deaza-deoxyadenosine substitutions on Coralyne Binding

We used the same duplex system **plusA8·minusA8** (named as **A·B** for simplicity) that we used previously in chapter 2, where eight deoxyadenosine nucleotides are placed across from each other in a duplex that is stabilized by seven Watson-Crick base pairing at both ends. We replaced different numbers of adenines at certain positions with 7-deaza-deoxyadenosine nucleotide (Figure 3.3), and investigate the stability of these duplexes in the absence and presence of coralyne by CD melts.

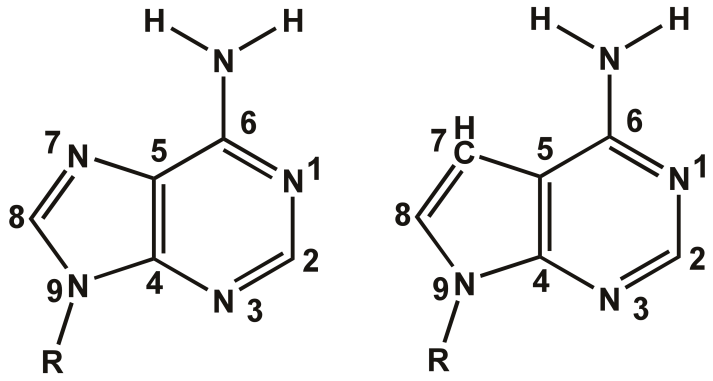

Figure 3.3: Chemical structures of deoxyadenosine (right) and 7-deaza-deoxyadeosine (left). R at 9 position represents the deoxyribose.

CD spectra of selected duplexes in the presence and absence of coralyne are given in Figure 3.4. In the absence of coralyne, changes were observed in the CD spectra of **(A·B)** duplex with 7-deaza-deoxyadenine substitutions. Substitution of only one nucleotide, **(A·B)** vs. **(1A·B)** resulted in very minor changes. However, increasing the

number of substitutions resulted in additional changes. The CD intensity at 220 nm and 247 nm decreased in intensity, while spectra were shifted to shorter wavelength at around 280 nm. These changes are most likely associated with the changes of the stacking interactions between adenines by the 7deaza-deoxadenosine substitutions. In the presence of the coralyne again there were changes associated with the increasing number of substitutions. Especially the intensity of the induced CD band between 300-360 nm decreased with the incorporation of multiple 7-deaza-deoxyadenosine instead of deoxyadenosine. This result indicates the reduced affinity of coralyne to the modified duplexes with increasing number of substitutions compared to the native **(A·B)** duplex**.** (Figure 3.4).

The complete list of melting temperatures is given in table 3.1. First, in the absence of coralyne the T_m of the all duplexes investigated in here changed by ± 1.5 ^oC, which is in the error range of our T_m measurement. On the contrary, there was a clear $_{Tm}$ decrease in the modified duplexes in the presence of coralyne. The T_m of the duplex decreased from 63 ^oC **(A·B)** to 59.5 ^oC **(1A·B)** and 53.5 ^oC **(3A·B)** by the substitution of only one and four adenosines with 7-deaza-deoxyadenosine respectively. Further destabilization up to T_m of 45.5 ^oC is observed by the incorporation of additional 7-deaza-deoxyadenosine across the already modified sequence **(3A·2B)**.

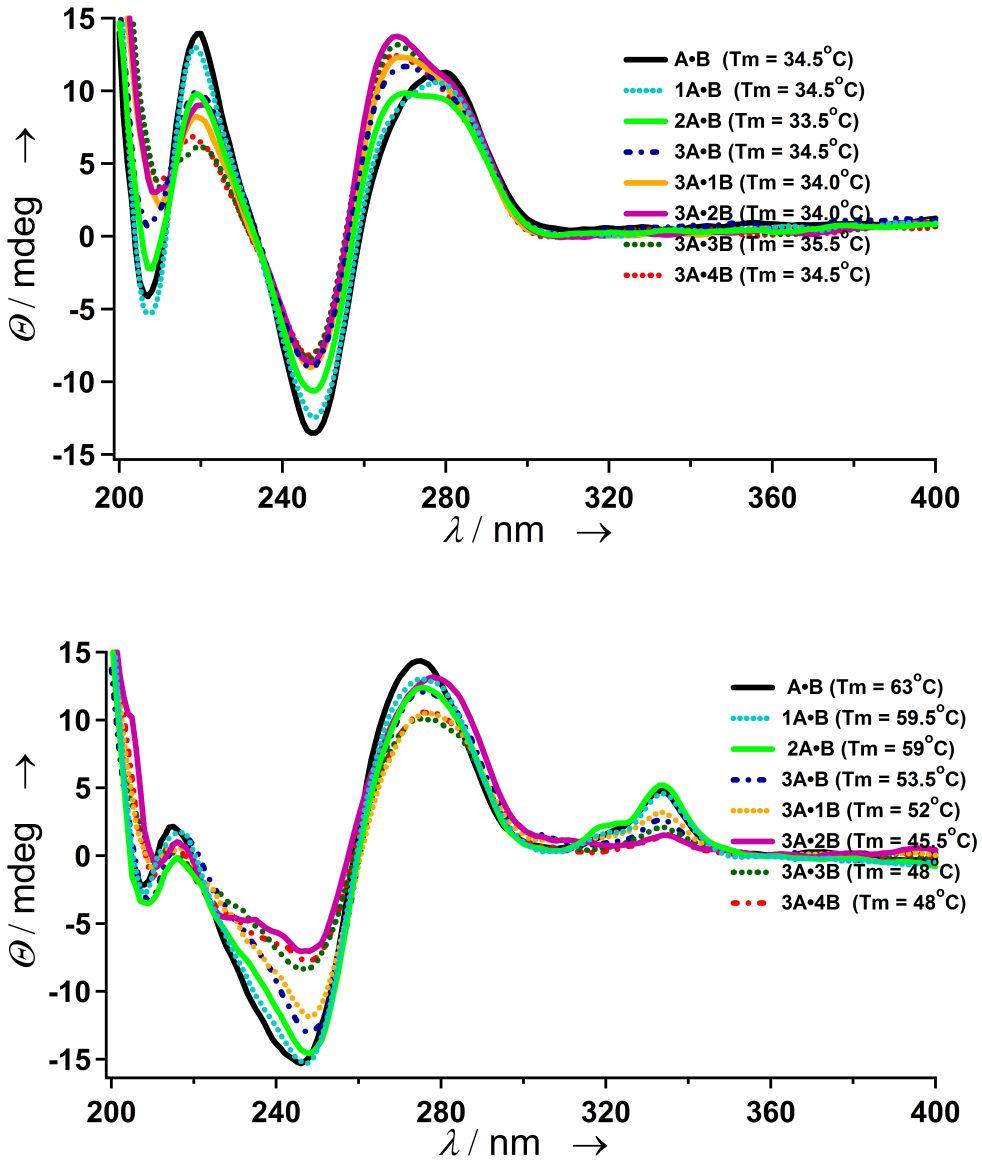

Figure 3.4: CD spectra of **(A·B)** duplexes in the absence (top panel) and presence (bottom panel) of coralyne at 5 °C.

Table 3.1: Melting temperature of all the sequences in the presence and the absence of coralyne. **7** refers to the 7-deaza-deoxyadenosine nucleotide. All of the sequences have the same flanking Watson-Crick base paired region represented by the lines in the table. For instance:

in **(A·B) duplex ;** **A - 5’-** GAC CCG C AA AAA AAA CCT CGC C **-3’**
B - 3’- CTG GGC G AA AAA AAA GGA GCG G **-5’**

Name	Sequence	T_m(^{o}C) without coralyne	T_m(^{o}C) with coralyne
(A·B)	5’-_____AAAAAAAA_____-3’ 3’-_____AAAAAAAA_____-5’	34.5	63.0
(1A·B)	5’-_____AAAA**7**AAA_____-3’ 3’-_____AAAAAAAA_____-5’	34.5	59.5
(A·1B)	5’-_____AAAAAAAA_____-3’ 3’-_____AAAA**7**AAA_____-5’	33.5	59.0
(2A·B)	5’-_____AA**7**A**7**AAA_____-3’ 3’-_____AAAAAAAA_____-5’	33.5	59.00
(A·2B)	5’-_____AAAAAAAA_____-3’ 3’-_____AA**7**A**7**AAA_____-5’	32.5	59.5
(A·3B)	5’-_____AAAAAAAA_____-3’ 3’-_____AAA**7**A**7**AA_____-5	33.5	58.0
(A·4B)	5’-_____AAAAAAAA_____-3’ 3’-_____AAA**77**AAA_____-5’	34.5	58.5
(1A·1B)	5’-_____AAAA**7**AAA_____-3’ 3’-_____AAAA**7**AAA_____-5	33.5	59.00
(1A·2B)	5’-_____AAAA**7**AAA_____-3’ 3’-_____AA**7**A**7**AAA_____-5’	33.0	55.0
(1A·3B)	5’-_____AAAA**7**AAA_____-3’ 3’-_____AAA**7**A**7**AA_____-5’	33.0	53.0
(1A·4B)	5’-_____AAAA**7**AAA_____-3’ 3’-_____AAA**77**AAA_____-5’	34.5	51.50
(2A·1B)	5’-_____AA**7**A**7**AAA_____-3’ 3’-_____AAAA**7**AAA_____-5’	33.0	53.5

Table 3.1 (continued):

Name	Sequence	T_m(^{o}C) without coralyne	T_m(^{o}C) with coralyne
(2A·2B)	5’-_____AA**7**A**7**AAA_____-3’ 3’-_____AA**7**A**7**AAA_____-5’	34.0	50.0
(2A·2B)	5’-_____AA**7**A**7**AAA_____-3’ 3’-_____AA**7**A**7**AAA_____-5’	34.0	50.0
(2A·3B)	5’-_____AA**7**A**7**AAA_____-3’ 3’-_____AAA**7**A**7**AA_____-5’	34.0	50.0
(2A·4B)	5’-_____AA**7**A**7**AAA_____-3’ 3’-_____AAA**77**AAA_____-5’	34.0	50.5
(3A·B)	5’-_____AA**7777**AA_____-3’ 3’-_____AAAAAAAA_____-5’	34.5	53.5
(3A·1B)	5’-_____AA**7777**AA_____-3’ 3’-_____AAAA**7**AAA_____-5’	34.0	52.0
(3A·2B)	5’-_____AA**7777**AA_____-3’ 3’-_____AA**7**A**7**AAA_____-5’	34.0	45.5
(3A·3B)	5’-_____AA**7777**AA_____-3’ 3’-_____AAA**7**A**7**AA_____-5’	35.5	48.0
(3A·4B)	5’-_____AA**7777**AA_____-3’ 3’-_____AAA**77**AAA_____-5’	34.5	48.0

3.3.3 Plausible Base Pairing

The obvious destabilization of these duplexes by the incorporation of 7-deaza-deoxyadenosine clearly indicates the importance of the N7 position in the A·A base pairing. Just taking the importance of N7 in A·A base pairing into consideration, one can eliminate structure 1c-1g where N7 do not involved in the H-bonding. Besides, taken together with the importance of protonation of N1 in A·A base pairing and the duplex

being anti-parallel, one can clearly point out 1a as the feasible base pairing. This base pairing is asymmetric where both N1 and N7 are involved in the base pairing with N6-amino group of the adenine in the opposite strand. In this base pairing both of the adenines adopt anti conformation. While one adenine is using the Watson-Crick face, the other is using Hoogsteen face of the base in order to participate in the base pairing.

We rationalize that such a base pairing can also explain the variation of the T_m in the modified 7-deaza-deoxyadenosine sequences. As demonstrated in table 3.1, the decrease in T_m of these duplexes is not proportional with the number of modifications; rather the place of the modification plays a role. For instance, both **(1A·B)** and **(2A·B)** duplexes have the same T_m of 59 ^{o}C even though they have one and two 7-deaza-deoxyadenosine substitutions respectively. Again the T_m of **(1A·2B)**, **(1A·3B)**, **(1A·4B)** and **(2A·1B)** vary between 51.5 ^{o}C and 55 ^{o}C event though they all have three substitutions. This analogy can be extended to duplexes in other series which have four or more substitutions. We rationalized that in this asymmetric base pairing (1a), two scenarios can be possible as seen in Figure 3.5. In the first case scenario, adenines can align on one strand as using only the Watson-Crick face and the other strand using only the Hoogsteen face (Figure 3.5a). In the second case scenario, they can alternate in the same strand as using the Watson-Crick face or Hoogsteen face (Figure 3.5b).

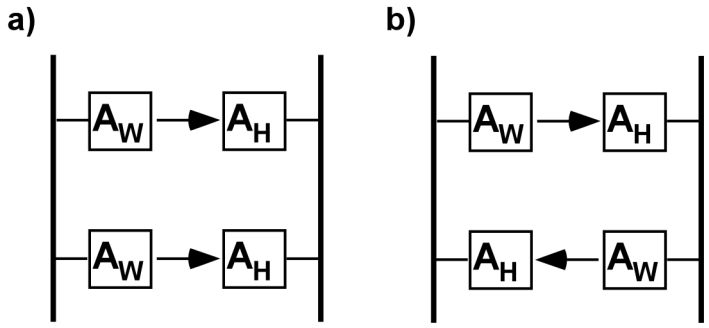

Figure 3.5: Possible alignment of adenines at the opposite strands to satisfy the possible base pairing 1b. A_W and A_H represent adenine nucleotides involved in base pairing by using their Watson-Crick face and Hoogsteen face respectively.

It is certain that the disruption of the base pairing by 7-deaza substitution at certain site, will affect the coralyne binding in this specific site. This specific site might stay empty without disturbing the other coralynes bounded to the neighboring sites, or the binding of coralyne to the whole duplex can be rearranged in order to maximize the filled sites. If there is no arrangement, and the sites were left empty, then whether it is the first scenario (Figure 3.5a) or the second (Figure 3.5b), the stability of the duplex should decrease proportionally to the number of 7-deaza substitutions. If there is rearrangement still the T_m should decrease proportionally with the number of 7-deaza substitutions if the adenines are arranged according to the first scenario (Figure 3.5a). However, if the second scenario is right where the adenines can also be rearranged to maximize the coralyne binding sites, then one can expect not only the number of substitutions but the place of the substitution might play a critical role in duplex stability. Based on where coralyne is binding, the adenines can alternate between using Watson-Crick face and Hoogsteen face in order to participate in the base pairing. In other words a certain adenine which is using the Watson-Crick face initially can switch to use its Hoogsteen

face in order to maximize H-bonding and stacking interactions if the 7-deaza-deoxyadenosine modification is in close proximity and altering the binding site of coralyne. In this respect, then coralyne is behaving like a molecular switch where the preferential binding of the coralyne results in the realignment of adenosines.

3.4 CONCLUDING REMARKS

In conclusion, here we have presented evidence for the feasible base pairing of A·A induced by the binding of coralyne at neutral pH. We have previously shown the unique specific recognition of adenine oligos and the induced secondary structure change by coralyne [65]. Coralyne can assemble adenine oligos into anti-parallel duplex at neutral pH which is different from the poly(rA) structure formed under acidic conditions [60]. In here taken consideration into the importance of protonation of N1, we have shown that N7 is also involved in the base pairing. Our T_m results support the formation of the asymmetric base pairing between two adenines as shown in 1a. This base pairing is asymmetric and stabilized by two H-bonds where both the N1 and N7 are making a single H-bond to N6-amino of the opposite strand. Furthermore, we have provided evidence that not only the number of substitution but the place of these substitutions is rather important resulting in a possible dynamic base pairing. Dynamic base pairing formation signifies once more the importance of the small molecule in recognition of the nucleic acid structure. In here we believe, coralyne binding not only promotes the base pairing between two adenines unless otherwise do not exist but also dictates how adenines will involve in the base pairing, which face will they use in order to participate

in the base pairing, Watson-Crick or Hoogsteen. Efforts are underway in our laboratory to determine the A·A base pairing by x-ray crystallography.

CHAPTER 4

SMALL MOLECULE RECOGNITION OF POLYADENYLIC ACID

4.1 INTRODUCTION

The interaction of small molecules with nucleic acids has been a subject of constant investigation for several decades [19, 28, 44, 69-71]. There is considerable interest in understanding these interactions at a level of detail that would allow the design of molecules that bind tightly and selectively to specific nucleic acid structures. In this respect, the polymorphism of DNA and RNA provide an opportunity for new therapeutic targets *in vivo*. Thus, the binding of small molecules to non-canonical nucleic acid structures (i.e. non-B-form and non-A-form structures), such as triplex and quadruplex DNA, has been a major focus of rational drug design [28]. Additionally, nucleic acids have emerged as the material of choice for the creation of supramolecular assemblies in nanotechnology. As we have discussed previously, small molecule-DNA interactions represent another dimension for expanding the functionality and dynamics of DNA-based nanotechnology [53, 56, 57, 72, 73].

Among the single stranded nucleic acids, poly(A) is of particular biological relevance due to its role in gene expression. Poly(A) tails, synthesized by poly(A) polymerase (PAP), are found at the 3′ end of mRNAs. These polyadenylate tails are important factors that contribute to the stability and maturation of mRNA, as well as the initiation of transcription [74, 75]. Since the discovery of over expression of Neo-PAP, a human PAP, in human cancer cells in comparison to normal or virally transformed cells

[76, 77], it has been suggested that the poly(A) tails of mRNA may represent a tumor specific target [78].

At present, only four molecules have been reported that bind poly(A) with high affinity and good selectivity over Watson-Crick duplexes (Figure 4.1) [78-81]. These four molecules have similar overall structures. Each has four fused six membered rings, which gives rise to a crescent shape and a larger surface area than the planar molecules typically found to intercalate DNA and RNA duplexes (e.g. the acridines and phenanthridines). Among the four reported poly(A) binders, coralyne and palmatine have dimethoxy groups attached at different positions, while berberine and sanguinarine have methylene dioxy groups that form five membered rings at the ends of the extended ring systems. Only two of these compounds, coralyne and sanguinarine, have been demonstrated to induce the formation of a poly(A) self-structure upon binding [78, 80].

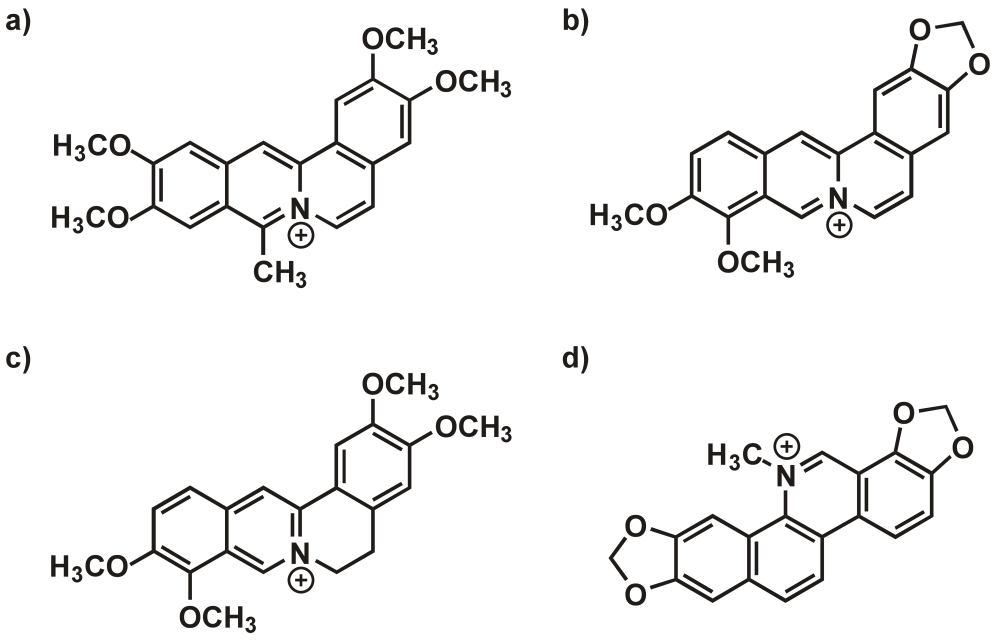

Figure 4.1: Chemical structures of (a) coralyne (b) berberine (c) palmatine (d) sanguinarine.

The poly(A) self-structure formed by in the presence of coralyne and sanguinarine has been suggested to be very similar to the coralyne-induced self-structure of poly(dA), a structure that is most stable around neutral pH [78]. As discussed in chapter 2, we initially demonstrated that the poly(dA) self-structure is an anti-parallel duplex with presumably A·A base pairs that is distinct from an earlier reported parallel stranded poly(A) duplex that is hemi-protonated and most stable under slightly acidic conditions [65].

In this chapter, we report an investigation into the importance of molecular structure for ligands binding to poly(A). We provide evidence that size and a crescent shape are important features for small molecules that induce and stabilize the poly(A) self-structure. Our studies demonstrate that relatively small changes in molecule structure result in tremendous changes in binding affinity for poly(A). Furthermore, our investigations reveal that the behavior of small molecule-driven nucleic acid assembly of the poly(A) self-structure is such that classical approaches for analyzing ligand-nucleic acid interactions can over- or under-estimate the binding constants of some molecules for poly(A). In particular, we report that that association constant of coralyne for poly(A) is higher than previously appreciated and among the highest reported for any small, planar molecule that binds to a nucleic acid double helix. Our analysis of the binding of molecules with lower affinity than coralyne demonstrate the multi-ligand dependence of the poly(A) self-structure, a system that mathematically is similar to the transition observed in micelle formation at the critical micelle concentration (CMC) [82-85]. The method used for making such measurements and that used for data analysis could prove

important for a wider range of nucleic acid systems in which a particular secondary structure is only stable upon binding of a particular small molecule.

4.2. EXPERIMENTAL PROCEDURES

4.2.1 Materials

Berberine chloride and coralyne chloride were purchased from Sigma and the concentrations of stock solutions in H_2O are determined spectrophotometrically by using previously published extinction coefficients; berberine $\varepsilon_{344} = 22500\ M^{-1}\ cm^{-1}$, coralyne $\varepsilon_{420} = 14500\ M^{-1}\ cm^{-1}$. Ellipticine was purchased from EMD Biosciences Inc and the stock solution was prepared in DMSO, ellipticine $\varepsilon_{295} = 60000\ M^{-1}\ cm^{-1}$. Heterocyclic Azacyanines are synthesized as described in 4.2.2. Stocks are also prepared in DMSO and the concentrations are measured spectrophotometrically using previously reported extinction coefficients. Poly(A) was purchased from Sigma and the $\varepsilon_{258} = 9800\ M^{-1}cm^{-1}$ in base is used to calculate the concentration. All samples are prepared in 1X BPE (6 mM Na_2HPO_4, 2 mM NaH_2PO_4 and 1 mM Na_2EDTA) with 20 mM NaCl at pH 7.0, unless otherwise mentioned.

4.2.2 Synthesis and Characterization of Azacyanines

Azacyanines were synthesized as described by Huang et al. and characterized by 1D (^1H, ^{13}C) and 2D (HSQC, HMBC) NMR spectroscopy, accurate mass spectrometry, and elemental analysis. Aza3 was also characterized by X-Ray crystallography (crystallization was effected from a saturated methanol-water solution at RT). Extinction coefficients were determined by UV-Vis spectrophotometry, using at least 5 different samples with different concentrations of the compound.

7*H*-1,13-Dimethyldibenzoimidazolo[1,2-*a*:2',1'-*d*][1,3,5]-triazin-6-ium (aza3). ^{1}H NMR (500 MHz, DMSO-*d*6): δ3.84, (s, 6H), 6.56 (s, 2H), 7.52 (m, 4H), 7.64 (m, 2H), 7.77 (m, 2H). ^{13}C NMR (125 MHz, DMSO-*d*6): δ29.58, 57.69, 110.89, 111.87, 125.08, 125.30, 128.19, 131.70, 149.68. HR-ESI: Mass observed 290.14066, mass calculated 290.14057 (-0.3 ppm). Elemental Analysis: $C_{17}H_{15.96}N_{4..96}I_1$ (calculated $C_{17}H_{16}N_5I_1$). Extinction coefficient in DMSO: ε_{343} = 47700 $M^{-1}cm^{-1}$. Yield 53%. The crystal coordinates are deposited to Cambridge Crystallographic Data Centre with deposition number 671506.

7*H*-Dibenzothiazolo[1,2-*a*:2',1'-*d*][1,3,5]triazin-6-ium (aza4). ^{1}H NMR (500 MHz, DMSO-*d*6): δ6.68 (s, 2H), 7.68 (s, 2H), 7.86 (s, 2H), 7.97 (s, 2H), 8.28 (s, 2H). ^{13}C NMR (125 MHz, DMSO-*d*6): δ61.40, 114.11, 124.83, 125.29, 127.65, 129.58, 139.97, 167.64. HR-ESI: Mass observed 296.03176, mass calculated 296.03162 (-0.5 ppm). Elemental Analysis: $C_{15}H_{10.66}N_{3.00}S_{2.02}I_{0.99}$ (calculated $C_{15}H_{10}N_3S_2I_1$). Extinction coefficient in DMSO ε_{387} = 32100 $M^{-1}cm^{-1}$. Yield 10%.

7*H*-3,11-Dimethoxydibenzothiazolo[1,2-*a*:2',1'-*d*][1,3,5]-triazin-6-ium (aza5). ^{1}H NMR (500 MHz, DMSO-*d*6): δ3.90 (s, 6H), 6.58 (s, 2H), 7.86 (dd, 2H), 7.46 (dd, 2H), 7.84 (dd, 2H), 7.88(dd, 2H). ^{13}C NMR (125 MHz, DMSO-*d*6): δ56.97, 61.40, 109.20, 114.90, 117.20, 126.30, 130.83, 158.92, 165.97. HR-ESI: Mass observed 356.05362, mass calculated 356.05340 (-0.2 ppm). Elemental Analysis: $C_{17}H_{13.91}N_{3.03}S_{2.04}O_{1.99}I_{0.99}$ (calculated $C_{17}H_{14}N_3O_2S_2I_1$). Extinction coefficient in DMSO: ε_{407} = 32000 $M^{-1}cm^{-1}$. Yield 23%.

4.2.3 Circular Dichroism (CD)

CD spectra were acquired on a JASCO J-810 CD spectropolarimeter equipped with Peltier temperature control unit. Spectra were acquired using either a 1 cm or 0.5 cm path length cell. CD melting profiles are acquired by increasing the temperature from 5 to 80 ^{o}C or 95 ^{o}C with 5 ^{o}C steps. All the samples were 60µM nucleotide base and 15 µM small molecule in 1X BPE buffer with 20 mM NaCl at pH 7.00. Job Plot analysis was carried out at 5^{o}C, with the combined concentration of aza and poly(A)/4 fixed at 15 µM. The CD intensity at 245 nm was normalized with respect to poly(A) concentration. For pH dependence the sample is prepared in 1X BPE with 20 mM NaCl at pH 7 and then the lower pH samples were prepared by the titration of 1 M HCl into this sample.

4.2.4 UV-Vis Spectroscopy

UV-Vis absorbance measurements were performed using a HP 8453 UV-Vis diode array spectrophotometer equipped with an Agilent 89090A Peltier temperature control unit. The melting profiles were acquired from 5 to 95 ^{o}C with a rate of <1 ^{o}C/min for all four traces (two heating and two cooling). For dilution experiments samples of poly(A) 200 µM in adenine base and 50 µM in small molecule concentration were prepared in 1X BPE with 20 mM NaCl at pH 7. These samples were diluted with 1X BPE, 20 mM NaCl buffer to the desired concentration and spectra was collected after each dilution, all at room temperature. For high concentrations, a 1 cm path length cell was used, and for low concentration (<10 µM in adenine) a 10 cm path length cell was used to increase signal to noise. In addition, extended integration times (10 sec vs 1 sec) as well as averaging of three spectra was used to further increase signal to noise. Percent of poly(A)–small molecule (coralyne, ellipticine, aza3 and aza5) complex formed at each

concentration was determined by performing a least-squares fit of the corresponding UV absorption spectrum as a weighted sum of two absorption spectra in the longer wavelength region (where only the small molecule absorbs), which were the spectrum of 200 μM poly(A) (in nucleotide base), 50 μM small molecule and a spectrum of free small molecule. For aza4 and berberine instead of using a two state model a change in absorbance in the longer wavelength region was monitored with dilution. For aza4 the ratio of average absorbance between 413-433 nm (representing the bound aza4) to average absorbance between 339 -359 nm (representing unbound aza4) were plotted with respect to concentration. The spectra were corrected for the background using the average absorbance between 437-457 nm. For berberine average absorbance from 458-478 nm relative to 398-418 nm were used after corrected for the background absorbance between 510-530 nm. For CMC modeling, equation 16 of Kegel et. al. [84] was used for the fraction of the material associated in multimolecular assemblies. The best fits are obtained when q, the number of small molecules need for the assembly, is set to six or larger. Briefly;

The self assembly of small molecules in micelle formation is described as:

$$qA_1 \rightarrow A_q$$

where A_1 is the individual molecules and A_q is the fully formed assembly.

$\ln K = \ln(X_{A_q}) / (X_{A_1}^{q})$ X defines the mole fraction

C* is the critical micelle concentration which is defined as the concentration of small molecules at which half of them is absorbed into assemblies $C^* = (qC_q)/C = 1/2$

C is the molar concentration of small molecules in solution therefore $C = C_1 + qC_q$

F is the fraction of small molecules absorbed in the assembly; $F = qC_q / C$ and is a function of C/C*

$C/C^* = (F^{1/(q-1)}) / (2(1-f)^{q/(q-1)})$

4.3. RESULTS

4.3.1 Coralyne Binding to Poly(A)

The binding of coralyne to a variety of nucleic acid sequences and structures has been previously reported [37, 43, 59, 62, 86]. It is known that coralyne binds preferentially to DNA triplexes over duplexes, most likely due to the closer match between the shape of coralyne and the shape of Py-Pu-Py base triplets, compared to Watson-Crick base pairs. This form of selective binding, associated with a common shape between ligand and base pairing structure, has long suggested that coralyne binds to DNA triplexes through intercalation [86, 87]. A competition dialysis study first established, unexpectedly, that coralyne also binds poly(dA) more favorably than Watson-Crick DNA [62]. Subsequently, our laboratory determined that the coralyne-induced self-structure of poly(dA) is an anti-parallel duplex with a helix that can be incorporated between flanking Watson-Crick duplexes [65]. More recently, Xing et al. reported that coralyne binds to poly(A) with a K_a of 1.8×10^6 M^{-1} at pH 7 [78]. Here we reexamine the binding of coralyne to poly(A) under similar conditions, but with a different experimental approach.

In Figure 4.2a the long wavelength region is presented from UV-Vis absorbance spectra of coralyne in the presence and absence of poly(A). A red shift of ca. 12 nm is observed upon coralyne binding to poly(A), which is consistent with the π-π stacking of

coralyne with the adenine bases, such as occurs upon intercalation. The addition of coralyne to poly(A) also resulted in the formation of induced CD bands between 290-360 nm (Figure 4.2c). These induced CD bands confirm the binding of coralyne in the chiral environment of poly(A). Thermal denaturation, monitored by CD, revealed a T_m of 60 ^{o}C for the poly(A)-coralyne complex , which is consistent with the results of Xing et al. [78].

Dilution experiments, in which buffer is added incrementally to a sample containing a constant ratio of one coralyne molecule per four nucleotides of poly(A), were conducted as a means to obtain the binding constant of coralyne to poly(A). We have found that maintaining a constant ratio of coralyne:poly(A) by this approach facilitates determination of the association constant for the primary mode of coralyne binding to poly(A). Previously, it has been shown that this highest affinity mode of coralyne binding to poly(A) is consistent with binding through the intercalation of A·A base pairs, as the maximum level of binding for the highest affinity mode obeys the nearest neighbor exclusion; one ligand per two base pairs (or four adenine bases in the case of the poly(A) duplex self-structure) [78].

We have determined that the poly(A) self structure can bind additional coralyne molecules by what is apparently a different and lower affinity mode of binding (e.g. non-specific outside helical stacking). Samples containing higher coralyne:poly(A) ratios can contain coralyne molecules associated with poly(A) in a mixture of binding modes, which complicates association constant determination. We have also determined that working with lower coralyne:poly(A) ratios than one coralyne per four nucleotides complicates binding studies because the poly(A) self-structure is only present when a critical number of small molecules are bound to the RNA polymer. Thus, the more

standard approach of association constant determination by measuring a physical observable (e.g. fluorescence) while titrating a small molecule into an excess of nucleic acid binding sites can give spurious results if the nucleic acid is not initially in the state that exhibits the binding mode of interest. This latter point is clearly illustrated by molecules that bind poly(A) with lower affinities for poly(A) than coralyne (vide infra).

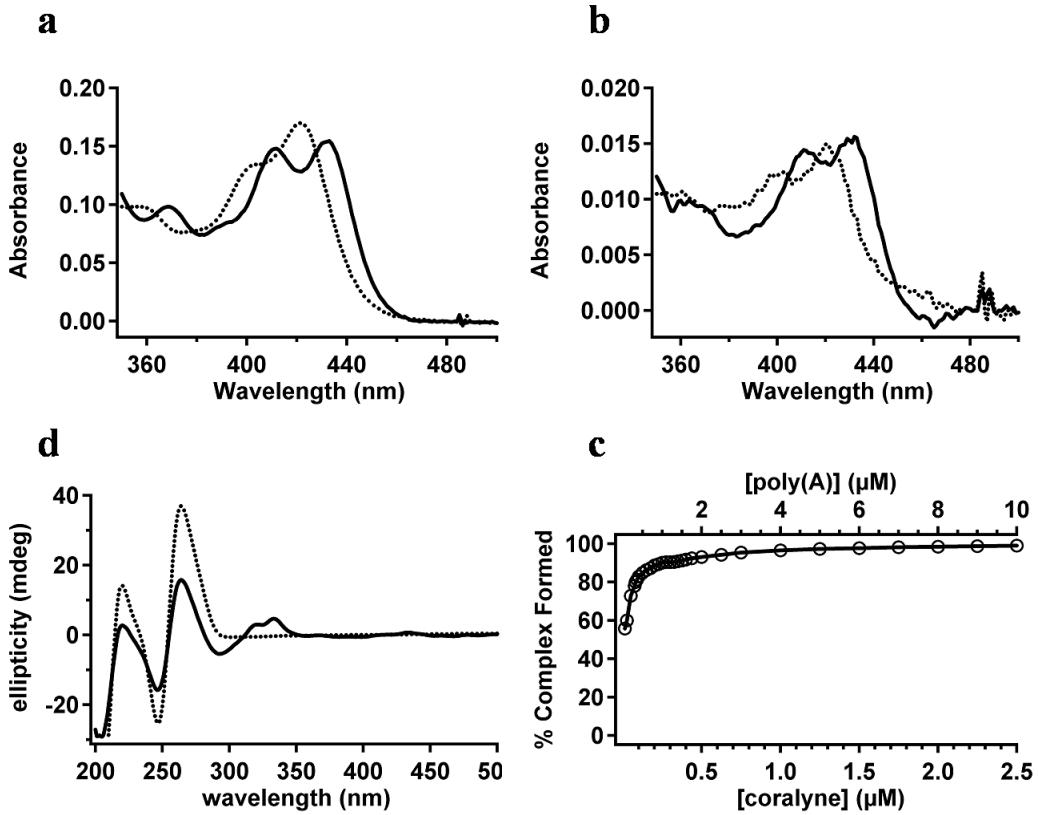

Figure 4.2: (a) The longest wavelength absorption region from UV-Vis spectra of samples containing 1.0 μM coralyne without (dashed line) and with (solid line) poly(A), 4 μM in nucleotide, at 22 °C. (b) The longest wavelength absorption region of the UV-Vis spectrum of 0.1 μM coralyne without (dashed line) and with (solid line) poly(A), 0.4 μM in nucleotide, at 22 °C. (c) CD spectrum of poly(A), 60 μM in nucleotide, in the absence and presence of 15 μM coralyne. Spectra collected at 20 °C. (d) Plot of the percentage of the poly(A)-coralyne complex present as a function of poly(A) and coralyne concentration. Data collected at 22 °C. For additional experimental details, see Materials and Methods.

For the coralyne-poly(A) dilution experiments a sample initially 200 μM in poly(A) nucleotide and 50 μM coralyne was diluted incrementally with buffer while UV-Vis absorption spectra were collected after each dilution. This procedure was continued until the sample was too dilute for spectra to be obtained with a satisfactory signal to noise ratio. The final spectra of these dilution studies were measured using a 10 cm path length cell and extended signal integration and spectrum averaging. In the case of coralyne, the sample diluted to a poly(A) concentration of 0.4 μM poly(A), 0.1 μM coralyne was near the practical limit for our instrumentation (Figure 4.2b). The percent of the poly(A)–coralyne complex formed at each concentration, starting with the highest concentration sample, was determined by fitting each dilution series spectrum with a two-state molecule. Briefly, least-squares fits were performed for the 380-480 nm region (a coralyne-only absorption region) of each UV-Vis absorption spectrum of the dilution study by a function that was a weighted sum of two absorption spectra; the spectrum of the initial 200 μM poly(A), 50 μM coralyne sample and the spectrum of a 50 μM coralyne sample with no added poly(A) (i.e. free coralyne). The UV-Vis spectrum of this highest concentration poly(A)-coralyne sample was taken to represent a spectrum corresponding to 100% poly(A)-bound coralyne, while the coralyne spectrum obtained in the absence of poly(A) was taken to represent 0% bound coralyne (i.e. no coralyne-poly(A) assembly).

The dilution curve obtained for bound coralyne using our two-state model is shown in Figure 4.2d. A plateau region exists from the highest sample concentration (i.e. 50 μM coralyne) down to 0.25 μM coralyne. All spectra were well fit by the two-state models with >95% bound coralyne. For samples diluted below 0.25 μM coralyne (1 μM

poly(A)), the reduced association of coralyne with poly(A) become apparent (i.e. the disassembly of poly(A) self-structure). However, even in the case of a sample diluted to a coralyne concentration of 0.1 μM, approximately 80% of coralyne is still bound to poly(A). A comparison of the spectra for 0.1 μM free coralyne and 0.1 μM coralyne with 0.4 μM poly(A) (Figure 4.2b) with corresponding spectra for 1 μM free coralyne and 1 μM coralyne with 4 μM poly(A) (Figure 4.2a) clearly illustrates that the association constant for coralyne binding to poly(A) is greater than 1×10^7 M^{-1}. The two-state model fits of these spectra, as described are provided in Figure 4.3. Dilution spectra obtained for concentrations below 0.1 μM coralyne, 0.4 μM poly(A) were also consistent with an apparent K_a of $>10^7$ M^{-1}, however, signal-to-noise rations for these samples was not sufficient to more precisely define the apparent association constant.

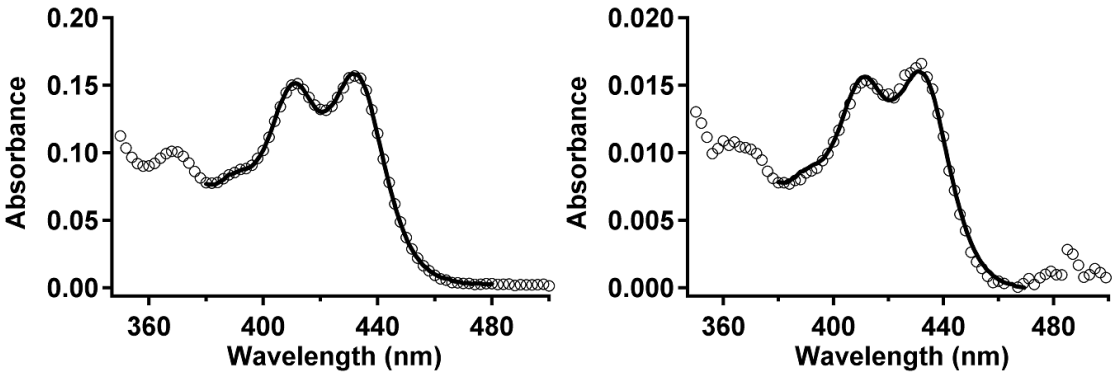

Figure 4.3: a) Longer wavelength region of the UV spectrum of poly(A)-coralyne complex which is 4 μM in adenine base and 1 μM in coralyne (markers). The solid line represents the best fit resulted from 96% bound and 4% unbound coralyne spectrum b) Longer wavelength region of the UV spectrum of poly(A)-coralyne complex which is 0.4 μM in adenine base and 0.1 μM in coralyne (markers). The solid line represents the best fit resulted from 82% bound and 18% unbound coralyne spectrum.

4.3.2 Ellipticine Binding to Poly(A)

We have previously demonstrated that proflavine (Figure 4.4), an intercalator of Watson-Crick duplexes with micromolar affinity, does not bind appreciably to polyd(A) (Chapter 2). Biver et al. have shown some level of proflavine binding poly(A), but through an outside stacking mode of binding that apparently does not induce the self-structure at neutral pH [88]. With the goal of discovering additional molecules that bind and stabilize the poly(A) self-structure, it was of interest to explore the structural requirements for molecules that bind poly(A) by investigating molecules intermediate to coralyne and proflavine in size and shape. Accordingly, we sought a molecule with a planar ring system that is only minimally larger than proflavine, but has a crescent shape. Ellipticine, a molecule also known to bind Watson-Crick base pairs that is similar in size to proflavine but with a crescent shape (Figure 4.4), was considered the best initial candidate for this a study.

The interactions of ellipticine and its derivatives with DNA have been studied extensively, due to their potential as anticancer drugs [89, 90]. The parent ellipticine molecule has been shown by X-ray crystallographic structure determination to bind duplex DNA by intercalation [91, 92] and, by fluorescence titration studies to bind calf thymus DNA with an apparent K_a of 10^7 M^{-1} (one of the highest association constants for a simple intercalator) [93]. Competition dialysis has also revealed that also ellipticine binds strongly to the hybrid duplex poly(A)·poly(dT), the DNA triplex poly(dA)·poly(dT)$_2$ and the DNA tetraplex formed by d($G_{10}T_4G_{10}$), but not poly(A) [30].

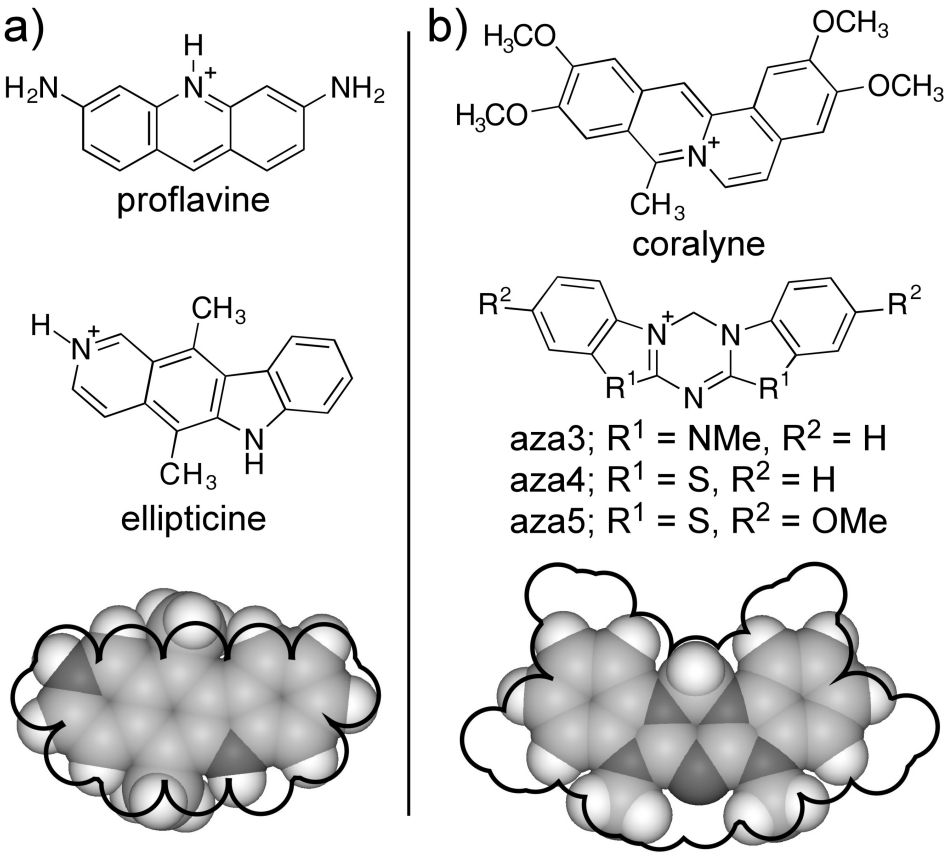

Figure 4.4: (a) Chemical structures of proflavine and ellipticine. Bottom of (a) is a space filling model of ellipticine overlaid with the outline of a space filling model of proflavine. (b) Chemical structures of coralyne and the azacyanine molecules. Bottom of (b) is a space filling model of aza3 overlaid with the outline of a space filling model of coralyne.

We first investigated the possible binding of ellipticine to poly(A) using both UV-Vis and CD spectroscopy. A sample containing one molar equivalent of ellipticine per four nucleotides of poly(A) exhibited a red shift in the longest wavelength absorption bands of ellipticine relative to a sample with the same ellipticine concentration but in the absence of poly(A) (Figure 4.5a). A weak induced CD band is also observed in the CD spectrum of the same sample (Figure 4.5b). Monitoring the change in ellipticity at 296 nm as a function of temperature revealed a T_m of 55 ^{o}C for the poly(A)-ellipticine

complex. A plot of the percentage of poly(A)-ellipticine complex formed as a function of ellipticine/poly(A) concentration, derived from a two-state model (Figure 4.5c), revealed an apparent association constant, K_a, of 2×10^5 M^{-1} (K_d of 4.6 μM) for ellipticine binding to poly(A).

Our discovery that ellipticine binds to poly(A), even though it is smaller than the previously identified molecules that bind the poly(A) self-structure (Figure 4.1), indicates that the crescent shape of ellipticine and coralyne is a critical feature for binding, as the more linear proflavine of similar size to ellipticine does not bind. The apparent association constant of ellipticine for poly(A), while comparable to proflavine and ethidium binding to Watson-Crick duplexes, is significantly less than that measured for coralyne binding to poly(A). Thus, it was of interest to then characterize the binding of molecules with a crescent shaped that are intermediate in size to ellipticine and coralyne.

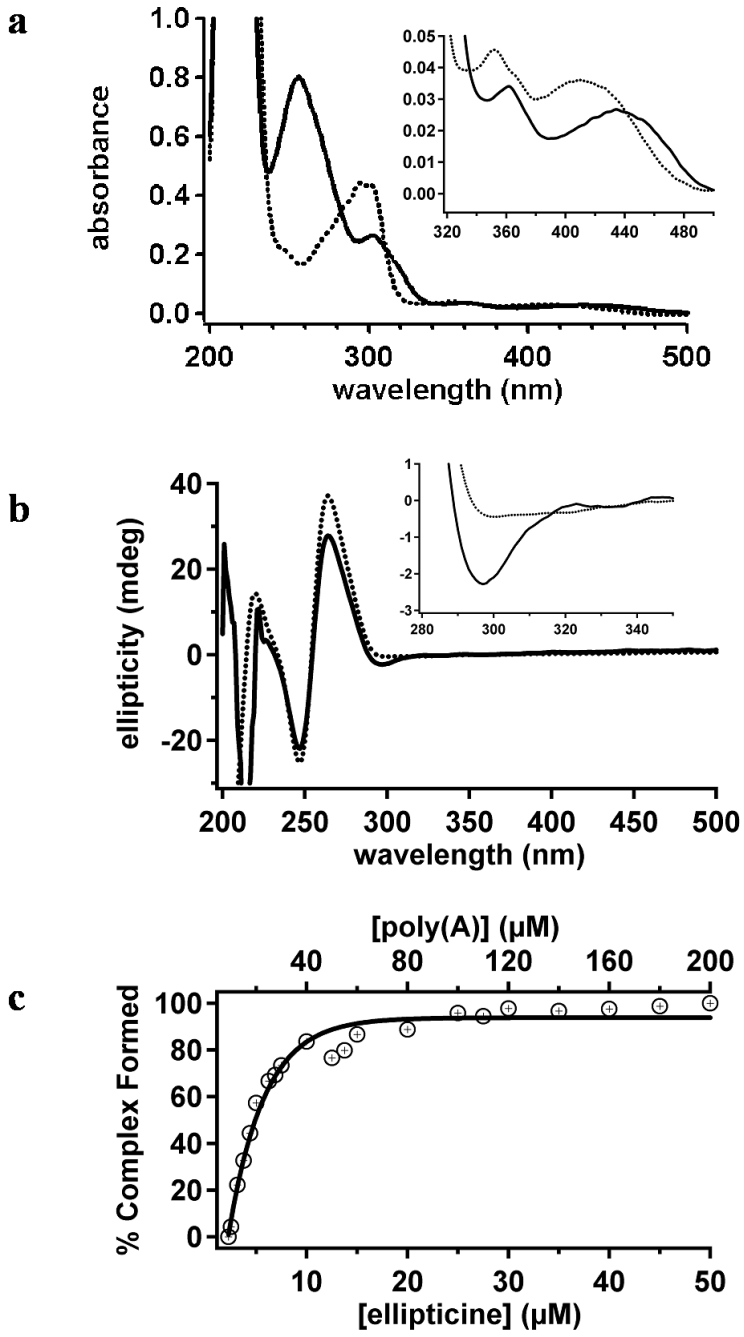

Figure 4.5: (a) UV spectrum of ellipticine in the absence (dashed line) and presence (solid line) of poly(A) at 22 ^{o}C. The inset shows the magnified longer wavelength region. (b) CD spectrum of poly(A) in the presence (solid line) and absence (dashed line) of ellipticine at 20 ^{o}C. The inset shows the magnified longer wavelength region. (c) Plot of percentage of poly(A)-ellipticine complex formed as a function of poly(A) and ellipticine (markers). Solid line represents the best fit with single exponential function.

4.3.3 Azacyanine Binding to Poly(A)

Azacyanines are a class of planar, bis-benzimidazole analogs that were originally prepared as putative ion channel blockers (Figure 4.4) [94-96]. These molecules satisfied our criteria of having crescent shaped and being intermediate in size to ellipticine and coralyne (Figure 4.4). The ring system of these molecules is very similar in size and shape to coralyne. Therefore, aza3 and aza4, which do not have exocyclic groups on the peripheral six membered rings, are smaller in overall surface area compared to coralyne (Figure 4.4), as well as the other known poly(A) ligands shown in Figure 4.1. The azacyanines are also of interest as potential poly(A) ligands because these molecules bind poorly to duplex DNA, with a K_{app} of around 10^3-10^4 M^{-1} (discussed in Chapter 5 in detail), which presents the possibility of azacyanines being selective poly(A) ligands.

UV-Vis and CD spectra of the azacyanines in the absence and in the presence of poly(A) are shown in Figure 4.6. The longest wavelength absorption bands in the UV-Vis spectrum of azacyanines are red shifted approximately 10 nm in the presence of poly(A), and a hypochromic effect of >25% is also observed (Figure 4.6). Both spectral changes indicate poly(A) binding with π-π stacking, which is suggestive of base intercalation [36]. The induced CD bands observed in the longer wavelength region also confirm the binding of azacyanines in the chiral environment of poly(A) (Figure 4.6).

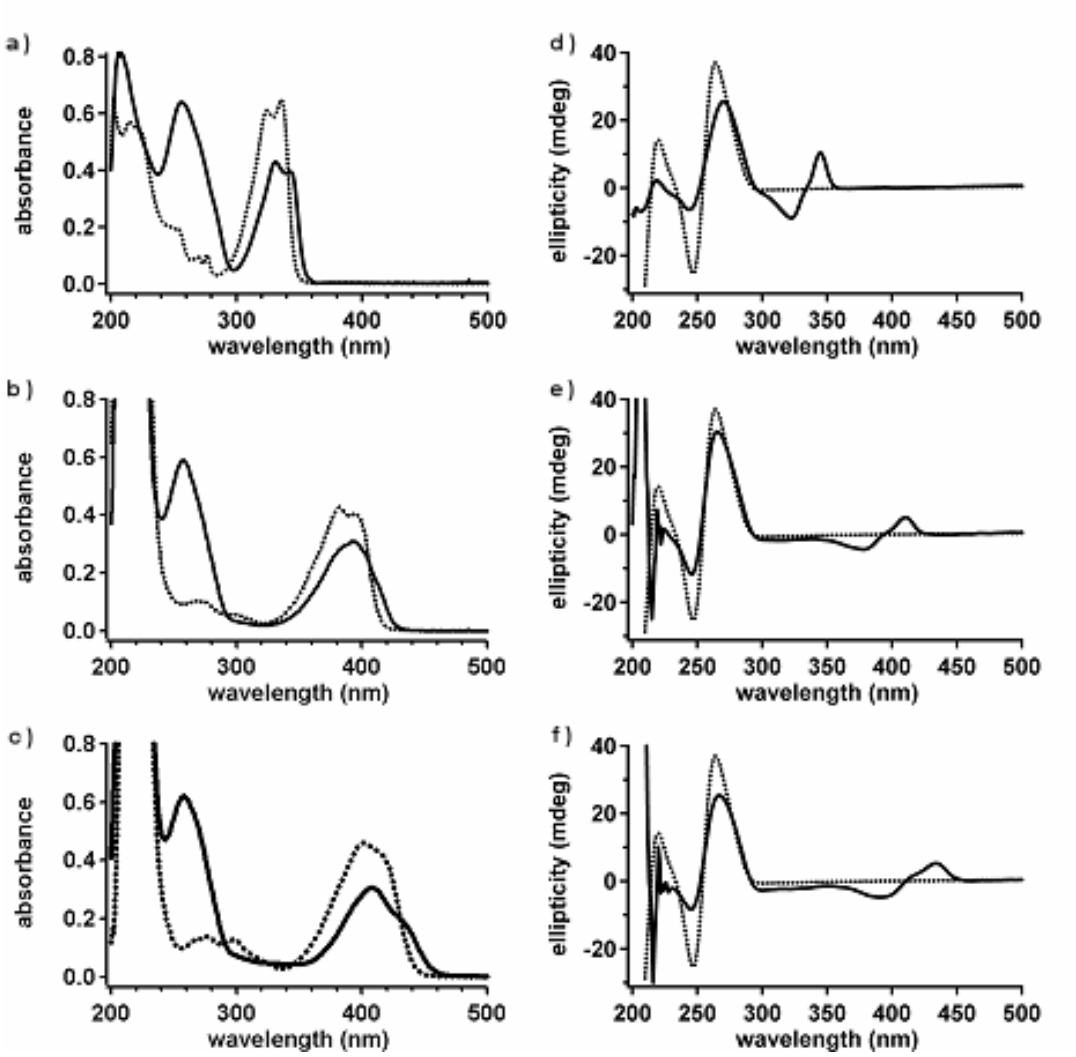

Figure 4.6: (a), (b) and (c) are UV-Vis spectra of aza3, aza4 and aza5 in the absence (dashed line) and in the presence (solid line) of poly(A) at 22 ^{o}C, respectively. (d), (e) and (f) are CD spectra of poly(A) without (dashed line) and with (solid line) aza3, aza4 and aza5 at 20 ^{o}C, respectively. All samples 60 µM in adenine base and 15 µM azacyanine are in 1X BPE with 20 mM NaCl at pH 7.0

CD spectra collected as function of temperature revealed cooperative melting transitions for poly(A) in the presence of aza3 and aza4 with transition midpoints, T_m, at 55 ^{o}C and 30 ^{o}C, respectively. These cooperative transitions, which are not observed for

poly(A) or the azacyanines alone, are consistent with the formation of an azacyanine-induced poly(A) self-structure. It was not possible to obtain an exact T_m for the poly(A) self-structure formed in the presence of aza5, as this molecule rapidly decomposed upon heating above 50 ^{o}C. Nevertheless, the poly(A)-aza5 complex heated to 40 ^{o}C maintained at least half the intensity of its maximally induced aza5 CD band, indicating a T_m of >40 ^{o}C. We note that aza4 also decomposed rapidly above 50 ^{o}C, but the T_m of the poly(A)-aza4 complex was sufficiently lower than this temperature that a T_m could be determined before appreciable decomposition.

To further investigate the nature of aza3 and aza5 binding to poly(A), job plot analyses were performed (Figure 4.7). These experiments revealed that azacyanines bind to poly(A) with a stoichiometry of one azacyanine per four adenine nucleotides. The same ligand: poly(A) stoichiometry that was previously determined for coralyne binding to poly(dA) and poly(A) [59, 78]. This binding ratio is again suggestive of intercalative binding, as the poly(A) self-structure is likely an antiparallel duplex, as is poly(dA), and the well-known nearest neighbor exclusion principle states that intercalators cannot bind at a ratio that is greater than one ligand per two base pairs (i.e. intercalated between every other base pair).

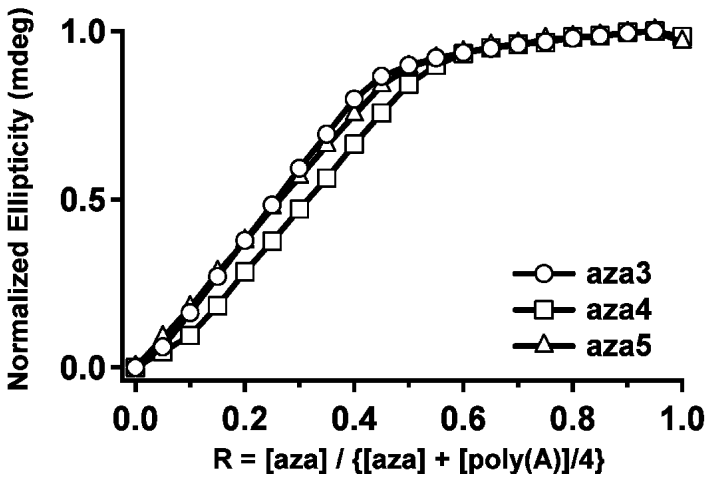

Figure 4.7: Job Plot analysis of poly(A) with azacyanines at 5 ^{o}C. CD intensity at 245 nm is normalized with respect to poly(A) concentration. The combined concentration of poly(A) and the azacyanine was kept constant at 15μM for each data point over the course of titration. The samples are in 1X BPE (6 mM Na_2HPO_4, 2 mM NaH_2PO_4 and 1 mM Na_2EDTA) with 20 mM NaCl at pH 7.0

Another distinguishing feature of the coralyne-induced poly(dA) self-structure is that it is destabilized as pH is lowered below neutrality [65]. This observation is significant given that poly(A) is known to form a parallel-stranded hemi-protonated self-structure that is most stable at pH < 4.5 [60, 97]. We have investigated the pH-dependent stability of the azacyanine-induced poly(A) self-structure using aza3 as a representative azacyanine ligand. Our pH titration experiments have confirmed that the aza3-induced poly(A) self-structure is destabilized as the sample pH is lowered below neutrality (Figure 4.8). Thus, the pH-dependent stability of the aza3-induced poly(A) self-structure again supports this structure as being the same as the coralyne-induced structure.

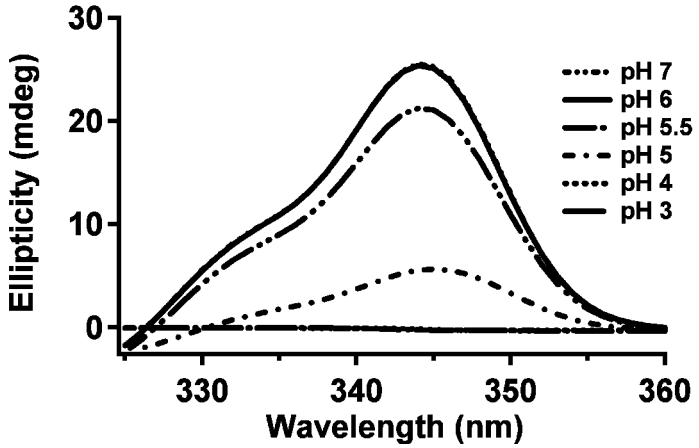

Figure 4.8: 325-360 nm region of the CD spectra of poly(A)-aza3 complex at varying pH conditions. Samples with pH values lower than 7 were prepared by the titration of a pH 7 sample with 1 M HCl. Samples were 60 μM nucleotide base, 15 μM coralyne, 20 mM NaCl and 1X BPE, at room temperature.

The dissociation of aza3 and aza5 from poly(A) was monitored by fitting UV-Vis spectra with a two-state model as described above for coralyne and ellipticine. The resulting curves revealed apparent association constants of 3.8×10^5 M^{-1} and 2.5×10^5 M^{-1} (K_ds of 2.6 μM and 4.0 μM) for aza3 and aza5, respectively (Figure 4.9). The binding constant of aza4 to poly(A) is discussed in the following section.

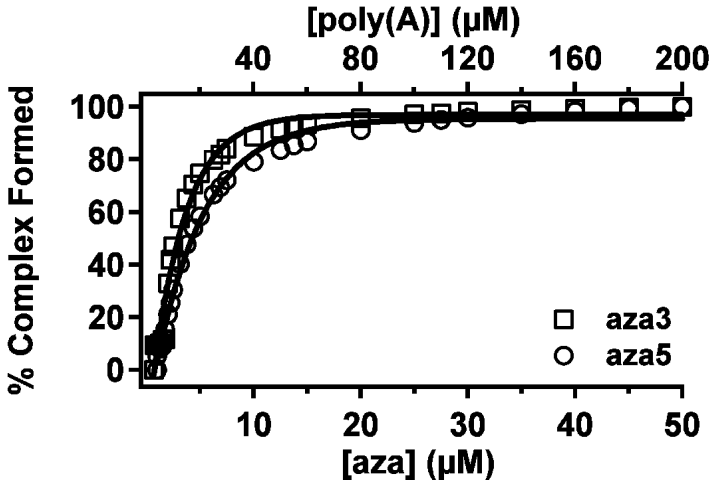

Figure 4.9: Plot of percentages of poly(A)-aza3 and poly(A)-aza5 complexes formed as a function of poly(A) and azacyanine concentration at 22 ^{o}C. Solid lines represent the best fit with a single exponential function.

4.3.4 Aza4 Binding to Poly(A)

The analysis used above to determine the association constants of several ligands to poly(A) utilized two-state models in which ligand-specific absorption bands from samples containing 50 μM ligand, 200 μM poly(A) were considered representative spectra for 100% ligand bound to poly(A). Support for these samples representing complete complex formation was based upon the observed plateau regions in the dilution curves, which persist down to at least a 10-fold dilution from the most concentrated spectra, and determination that T_m values for each of the ligand-poly(A) complexes are at least 15 °C above room temperature. Determination of the association constant of aza4 for poly(A) proved more challenging due to our lack of a sample in which aza4 was fully bound to poly(A). Our conclusion that all aza4-poly(A) samples investigated were only partially complexed was based upon the lack of a plateau region in the dilution curve, an

observation that was consistent with the T_m of the aza4-poly(A) complex being close to room temperature (i.e. 30 °C).

Despite the additional challenge encountered with determining the association constant of aza4 for poly(A), the dilution curve generated from the highest concentration aza4-poly(A) sample revealed an interesting and informative feature that was not apparent in the ligand-poly(A) dilution curves discussed above. In Figure 4.10 a dilution curve is shown for the aza4-poly(A) sample that was created by plotting the sample absorption ratio A_{423}/A_{349}. This first step of analysis includes no assumptions regarding whether the most concentrated spectrum represents fully complex aza4, only that the A_{423}/A_{349} absorption ratio is different for bound and unbound aza4. We note that, unlike the dilution curves presented above, the aza4-poly(A) dilution curve exhibits a cooperative transition between 0 and 25 µM aza4 (Figure 4.10).

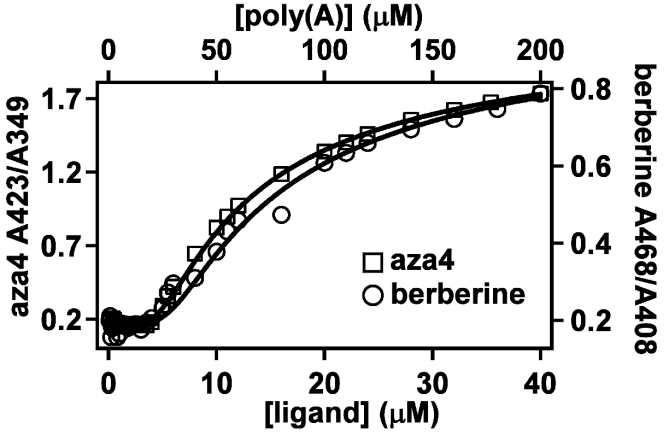

Figure 4.10: Plots of data obtained from dilution studies of aza4 (open squares) and berberine (open circles) in the presence of poly(A). Both curves are rms best fits of the CMC equation for six molecules being required for stabilization of the poly(A) self-structure. A sample of 200 µM in adenine base and 50 µM in small molecules containing 20 mM NaCl and 1X BPE is diluted with 1X BPE buffer containing 20 mM NaCl, spectra taken after each dilution at room temperature. For more information, see Materials and Methods.

The shape of the aza4-poly(A) dilution curve is reminiscent of curves that monitor that formation of multimolecular assemblies. Micelle formation is a classic example [82, 83, 85]. The cooperative transition observed in micelle formation as a function of surfactant concentration is explained by the self-association constant between surfactants and the number of surfactants needed to nucleate micelle formation. This analysis yields what is know as the *critical micelle concentration* or CMC, below which micelle stability decreases rapidly with surfactant concentration. The concept of the CMC for explaining the behavior of multimolecular assemblies has been successfully applied to a host of other molecular systems, such as the assembly of virus capsid proteins [84].

The aza4-poly(A) dilution data is fit well by the CMC equation using three independent parameters: the association constant of aza4 for poly(A), the number of aza4 molecules needed to stabilize the poly(A) self-structure and the fraction of complex formed at the highest aza4/poly(A) concentration. When all three parameters were allowed to vary the rms least squares best fit was obtained for an aza4-poly(A) association constant of 4.8×10^4 M^{-1}, a minimum stoichiometery of 6 aza4 and a level of assembly at 50 µM aza4/200 µM poly(A) of 75%. In the fitting procedure the minimum number of aza4 molecules required for assembly was restricted to being an integer number.

4.3.5 Berberine Binding to Poly(A)

The binding of berberine to poly(A) and polyd(A) has been previously studied by two laboratories. It has recently been reported that berberine binds to poly(A) with a K_a of 3.5×10^5 M^{-1} (at pH 7.1; 20 mM NaCl, 20°C) as determined by fluorescence titrations [81]. However, earlier competition dialysis experiments, with somewhat different buffer

conditions (pH 7; 185 mM NaCl, room temperature (20-22^{o}C)), revealed no appreciable binding of berberine to polyd(A) [62]. Given the fact that the similar shaped coralyne was able to bind tightly to both poly(A) and polyd(A) and induce a self-structure formation, the results of berberine binding to these nucleic acids reported by two different labs was puzzling. And, understanding the binding of berberine with nucleic acids is of particular interest, as this molecule exhibits antisecretory, antiinflammotaory, antibacterial, antimalarail and anticancer activities with low toxicity [98-101].

We have reinvestigated the binding of berberine to poly(A) using the same set of experiments presented above. UV-Vis and CD spectral changes are shown in Figure 4.11 for berberine in the absence and presence of poly(A). The absorption spectrum of berberine in the presence of poly(A) is red shifted and hypochromic in the longest wavelength region, and induced CD bands are observed between 300-390 nm as well. Monitoring the ellipticity change at 330 nm revealed a T_m of 30 ^{o}C. This T_m for the berberine-poly(A) is quite low compared to coralyne, which is of similar structure (Figure 4.1).

The dilution curve for berberine in the presence of poly(A) was also generated by plotting an absorption ratio rather than using a two-state model, like aza4, because there was no indication that the sample with the maximum concentration of 50 µM berberine and 200 µM poly(A) represented fully bound berberine. In the case of berberine, the dilution curve was generated using the absorption ratio A_{468}/A_{408} (Figure 4.10). The resulting data lacks a plateau region, consistent with the berberine-poly(A) complex T_m being near room, and exhibits a cooperative transition, indicating little assembly below approximately 5 µM berberine (MCA, minimum concentration required in order to

initiate assembly). Given the shape of this curve, the data was also fit by the CMC equation (Figure 4.10). An excellent fit was again obtained for the requirement that six berberine molecules be bound for stabilization of the poly(A) self-structure, with an association constant of 4×10^4 M^{-1} and 71.4% assembly at the highest concentration. For this sample, a slightly better rms fit was obtained for four berberine molecules being required for poly(A) assembly, with 3.2×10^4 M^{-1} and 64% assembled. In either case, it is clear that multiple berberine molecules are required for poly(A) self-structure assembly and that the association constant is less than 4×10^4 M^{-1} (i.e. $K_d > 25$ μM).

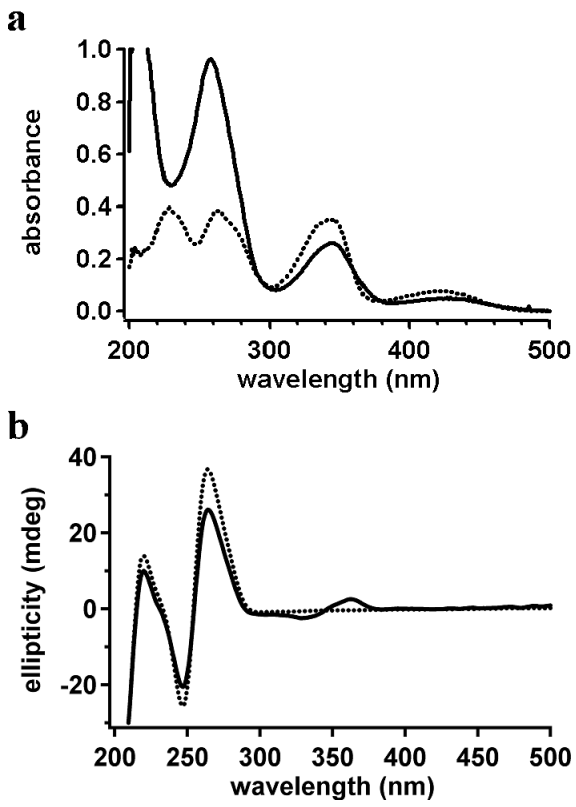

Figure 4.11: (a) UV spectra of berberine in the absence (dashed line) and in the presence (solid line) of poly(A) at 22 ^oC. (b) CD spectra of poly(A) without (dashed line) and with (solid line) berberine at 20 ^oC respectively. Samples were 60 μM nucleotide base, 15 μM berberine, 20 mM NaCl and 1X BPE.

4.4. DISCUSSION

We have used UV-Vis and CD spectroscopy to investigate the association of several small molecules with poly(A). These investigations have confirmed and expanded the set of ligands known to date that bind tightly to the poly(A) and induce self-structure formation at neutral pH. The longest wavelength absorption bands of the molecules investigated all exhibit a red shift and significant hypochromic effect upon binding to poly(A). These spectral changes are indicative of π-π stacking, which is consistent with base pair intercalation as the possible mode of binding [36].

The results of the experiments presented here illustrate the importance of the size and structure of molecules that induce formation of the poly(A) self-structure. Here we have demonstrated that coralyne binds poly(A) with an association constant that is among the highest observed for a planar molecule binding to a nucleic acid duplex (i.e. $K_a > 1\times10^7$ M^{-1}) and the highest reported to date for poly(A). In contrast, berberine, another alkaloid with a very similar shape, but with an extra ring system, showed greatly reduced binding (i.e. $K_a < 4\times10^4$ M^{-1}).

These observations motivated us to explore the possibility that molecules smaller than coralyne, but with a similar crescent shape, might also able to induce the formation of poly(A) self-structure. Ellipticine, a well-known intercalator of Watson-Crick duplexes was found to bind poly(A) with the respectable association constant of 2×10^5 M^{-1}. While interesting with regards to illustrating the molecular shape features that are important for poly(A) binding, the fact that ellipticine binds Watson-Crick duplexes with a similar affinity necessarily implies that this molecule is not selective for poly(A) over Watson-

Crick duplexes [89]. Consequently, we were motivated to find molecules in between ellipticine and coralyne that would bind poly(A), and possibly show selectivity.

Three azacyanines with the same six-five-six-five-six fused aromatic ring structure were selected for study due to their crescent shape and intermediate size between ellipticine and coralyne. All three azacyanines investigated were found to bind poly(A), with aza3 and aza5 exhibiting association constants that were comparable to that of ellipticine. Aza4 showed surprisingly reduced binding, similar to that of berberine. In this respect, azacyanines represent a promising new scaffold for the design of new molecules binding to poly(A) due the ease of their one-pot synthesis in addition to their lack of binding to duplex DNA [95, 96, 102]. One can easily modulate their binding by making different substitutions or additions of new functional groups to the ring system. All of the results demonstrated in here confirm that the size and the shape of the molecule is important in recognition of poly(A). In addition, comparison of coralyne and berberine, aza3 and aza4, aza4 and aza5 binding to poly(A) demonstrates the importance of functional groups in the small molecule structure for recognition (i.e. K_a of coralyne vs berberine, aza4 vs aza5).

Here we have also presented a simple yet effective methodology for obtaining binding constants for ligands of poly(A). For coralyne and berberine this approach has revealed association constants that differed from previously reported results obtained by more classical drug-DNA titrations. These studies indicate that coralyne binds to poly(A) with an association constant at least 5-fold greater, and berberine binds to poly(A) with an association constant of at least 100 fold lower, than what has been previously reported under similar conditions by more standard fluorescence titration measurements [78, 81].

Thus, the use dilution studies that take advantage of observed changes in UV spectra could prove more reliable in some cases than classical drug-DNA titrations experiments for determining association constants.

The binding of molecules to poly(A) is different from the binding of a molecule to a pre-structured nucleic acid, such as duplex DNA, where the binding of a molecule can be treated as a single event. In the case of poly(A), small molecule binding and secondary structure formation are coupled events as the poly(A) self-structure exits if and only if a small molecule is bound to the duplex. Therefore, if the free energy associated with binding of more than one ligand molecule is necessary to stabilize the poly(A) self-structure, then a cooperative transition should be expected. By diluting a concentrated poly(A)-small molecule complex, in which the poly(A) is initially in the double helical state, it is possible to capture the point at which the small molecules dissociates from the poly(A) self-structure and the self-structure itself is lost (Figure 4.12).

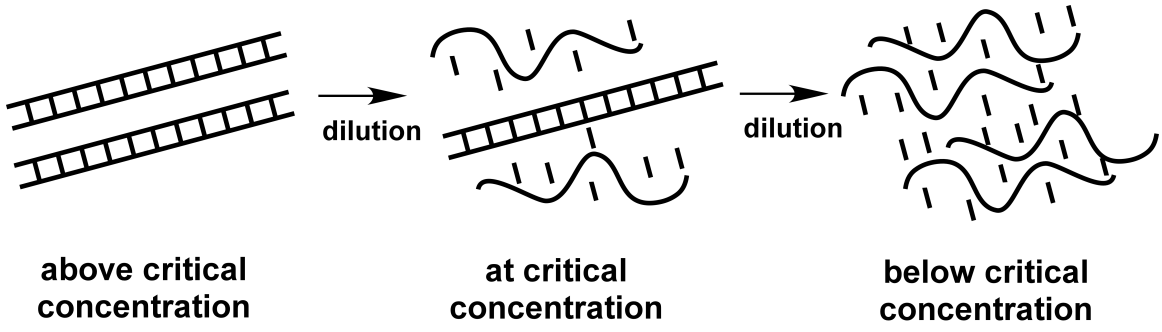

Figure 4.12: Disassembly of poly(A)-small molecule complex. Above a critical concentration (CMC) of small molecule poly(A) adopts anti-parallel double helical structure. By dilution of the sample small molecule starts to dissociate resulting in the dissociation of the strands until CMC. Further dilution below CMC results in the complete dissociation of the strands.

We propose that analysis of systems such as the poly(A)-coralyne assembly by classical drug-DNA binding assays may lead to ambiguous results because of the high cooperativity of small molecule binding. For example, in fluorescence titrations, generally a sample of free small molecule (most commonly 1 µM) is titrated with the nucleic acid. K_a or K_d, values are calculated by fitting the observed changes in the fluorescence spectroscopy using a two state model, i.e. the free and the bound state of the small molecule [103]. If the K_d of molecule is significantly higher than 1 µM ((e.g. coralyne, ellipticine, aza3 and aza5), then the poly(A) strands will be completely assembled to double helical structure at the first titration point. As we have demonstrated for molecules that bind tightly to poly(A) (e.g. coralyne, ellipticine, aza3 and aza5) the steepness of the slope of the binding curves are striking. There is a very narrow concentration range over which the small molecule and the nucleic acid assemble into a complex, and this feature is likely to be observed even if the fluorescence titration was initiated above the point of assembly. Titrations of the nucleic acid to high concentrations will show increased binding of the small molecule, which will yield an apparent association constant upon curve fitting. However, depending on the first titration point (and what percent is already assembled at this point) in the fluorescence titration, the subsequent titrations will only be able to establish binding curve partially compared to the binding curve obtained by dilution curves, leaving the initial assembly and the binding of the small molecule out.

In equilibrium dialysis experiments, nucleic acids samples in dialysis units are submerged into a solution of the small molecule of interest (commonly at a concentration of 1 µM) [62]. If the K_d of small molecule binding is lower than 1 µM (e.g. aza4 or

berberine binding to poly(A)) and if multiple molecules are needed to assemble the nucleic acid structure (MCA), there will be no apparent binding. Only small molecules that bind nucleic acids with a K_d of ≥ 1 μM will exhibit binding (e.g. coralyne to poly(A)). Ellipticine is an excellent example of a molecule that has a similar affinity for poly(A) as ethidium for duplex DNA, but does not show binding in an equilibrium dialysis assay [30], apparently because a particular concentration (or MCA) of ellipticine is needed before the assembly of poly(A) will be favored. Therefore, if the MCA of ligand binding to an unstructured nucleic acid is higher than 1 μM, then no binding by the ligand will be observed. The same argument might also explain the discrepancy in obtained association constants for berberine to poly(A) and polyd(A) by fluorescence titrations and the equilibrium dialysis, where MCA is established as 5μM by dilution studies [62, 81].

Then again, if multiple molecules are needed to favor poly(A) assembly, in fluorescence titration one would expect to see two stages under these conditions, K_d < 1μM. The first stage will be the lack time until reaching to the critical concentration followed by the second stage which is the assembly of poly(A) strands by binding of the small molecule. Consequently, fitting these two different events with a two state model may also result in ambiguous results [30, 62].

4.5. CONCLUDING REMARKS

In conclusion, there are only a few molecules know at present that bind poly(A) selectively and with high affinity. Here, by taking the size and the shape of the coralyne and proflavine in consideration, we have reported the binding of four additional

molecules that bind poly(A), ellipticine and heterocyclic azacyanines. Moreover, we have demonstrated that the size and the structure of the small molecule is extremely important in recognition of a particular nucleic acid structure. Additionally, we have demonstrated the utility of obtaining association constants for poly(A)-like systems by simply monitoring changes in UV spectra over the course of sample dilution. We propose that such dilution experiments are complementary to more classical drug-DNA binding experiments, and provide information that may be difficult to obtain by any other method.

We believe that similar simplistic rational design approaches can be used in order to increase the number of small molecules that bind to non-canonical nucleic acid structures with high affinity, and help us to understand the recognition of these structures by small molecules. In addition, we believe our studies with small molecule-driven assembly of nucleic acid structures suggest the possibility of using such associations for the creation of concentration-dependent nanoscale devices that respond sharply to changes in small molecule concentration (e.g. dilution or evaporation).

CHAPTER 5

INTERACTIONS OF AZACYANINES WITH HUMAN TELOMERIC DNA*

5.1. INTRODUCTION

It has been know for decades that GMP and poly(dG) are able to self associate and form gels in solution [104-106]. The basic structural motif for these gels has been identified as the G-quartet, formed by the arrangement of four planar Guanine bases held by Hoogsteen base pairing. G-rich DNA and RNA strands have been shown to form G-quadruplexes where two or more G-quartets stack on top of each other and form intramolecular or intermolecular four stranded structures (Figure 5.1) [9, 107-111].

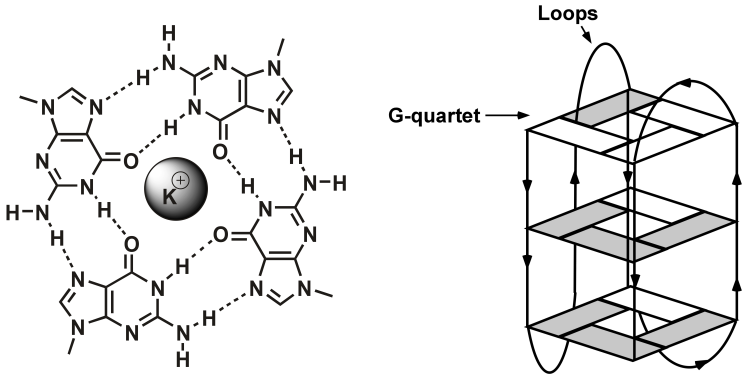

Figure 5.1: Structures of a G-quartet (G-tetrad) (left panel) and G-quadruplex formed by the stacking of three G-quartets. Each rectangle represents a guanine base.

---

* The majority of the work presented in this chapter will be published in Persil, Ö., et al. *Chembiochem*, **2008** (in publication) [102]. Copyright Wiley-VCH Verlag GmbH & Co. KGaA.Reproduced with permission.

G-quadruplex structures have been shown to be stabilized by the presence of monovalent cations, such as Na^+ or K^+ under physiological conditions [108, 112]. In the last decade, interest in these structures increased significantly due to the presence of G-rich sequences in telomeric DNA and promoter regions. These non-canonical DNA structures have become very promising as drug targets [8-12, 58, 113].

Telomeres, specialized repetitive DNA sequences, are found at the ends of eukaryotic chromosomes and are required for replication and stability. Human telomeric DNA is 5-8 kb and has a 100-200 nucleotide long single stranded overhangs at the 3' end that is composed of d(TTAGGG) repeats [114]. In normal somatic cells, after each round of cell division telomeres decrease in length, eventually resulting in apoptosis. However, in cancer cells telomeres do not shorten due to the activation of an enzyme called telomerase, a riboprotein complex. It has been shown that telomerase which is inactive normal somatic cells, is 80-85% active in human cancer cells [115]. At each replication cycle telomerase is able to lengthen the telomeres by using 3' single stranded overhangs as a template. The formation of G-quadruplex structures in these 3'single stranded regions has been shown to inhibit the activity of the telomerase [116-118]. Therefore, promoting G-quadruplex formation in these single stranded regions by small molecule binding has become an attractive approach for the development of anti-cancer drugs [8-12, 58, 113]. Small molecules with high affinity and high selectivity for G-quadruplexes have even begun to show medicinal promise, although the connection between G-quadruplex binding and *in vivo* activity may not always be obvious [119]. For example, quarfloxin, an antineoplastic targeting the rRNA-nucleolin complex, is presently in phase II clinical trials [27].

Two common strategies have been followed by most investigators seeking new ligands for G-quadruplex DNA. First, they have focused on heterocycles with a relatively large and planar surface area, which maximizes stacking with the ca. 1 nm^2 surface of a G-tetrad (e.g., TmPyP4, Figure 5.2) [120, 121]. Second, many have used multiple charges to increase electrostatic interactions with the high-charge density G-quadruplex (e.g., BRACO-19, Figure 5.2d) and to enhance the solubility of potential ligands (often necessary for ligands with large hydrophobic surfaces, e.g., TmPyP4) [122, 123]. While these strategies have resulted in several high affinity ligands for G-quadruplexes, most ligands do not exhibit high selectivity over duplex DNA.

Recently, several metallated ligands have been reported with substantial selectivity for G-quadruplex DNA [124]. As the authors have suggested, this may be, in part, due to interactions between the metal ion and the lone pairs of the carbonyl oxygen atoms of the exterior quartets, which are not accessible in duplex DNA. Higher-throughput, more combinatorial approaches have begun to show promise as well. For example, Balasubmaranian and coworkers have recently demonstrated dynamic combinatorial selection of quadruplex ligands via disulfide bond formation with a thiol-containing scaffold and side chains [125].

We previously discovered that a planar molecule slightly larger than a typical intercalator of Watson-Crick base pairs can selectively bind purine-purine base pairs [37, 59, 65]. Based upon this discovery, we hypothesized that planar, monocationic molecules that are marginally too large to intercalate duplex DNA, such as bispurine analogues, might selectively bind G-quadruplexes. In this chapter we report that the azacyanines (Figure 5.2) represent a new class of selective, submicromolar quadruplex ligands that

can be obtained by a facile synthetic route. The synthesis of azacyanines was previously reported by Kurth and coworkers; the route is one-pot and workup is by filtration [95, 96]. The route is general and succeeds for aminobenzimidazoles and aminobenzothiazoles. The synthetic ease makes the class amenable to library preparation for high-throughput screening.

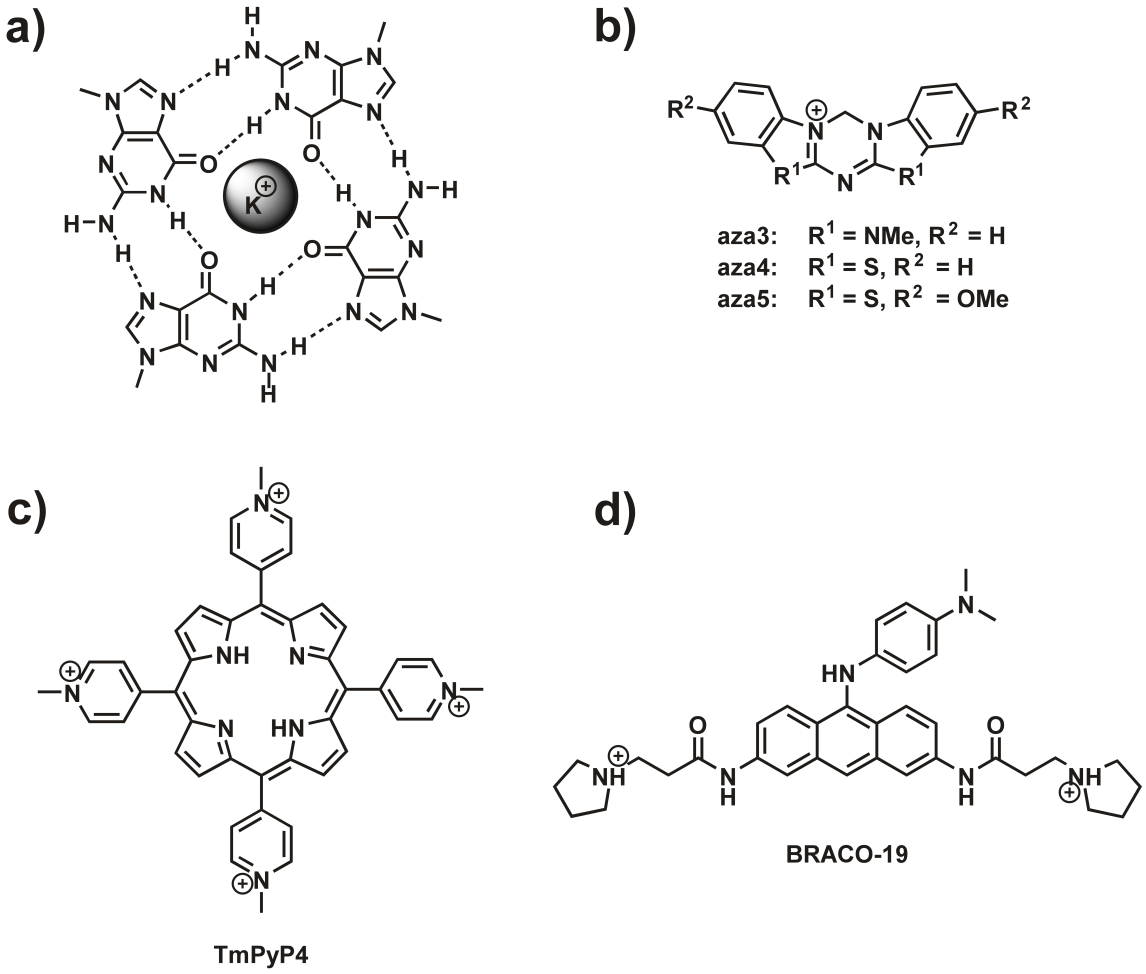

Figure 5.2: Structures of a) a G-tetrad and G-quadruplex ligands: b) the azacyanines, c) TmPyP4, and d) BRACO-19.

5.2. EXPERIMENTAL PROCEDURES

5.2.1 Materials

The oligonucleotide sequences d[TTGGG(TTAGGG)$_3$A] (tel24), d[AAAGGG(TTAGGG)$_3$AA] (tel26), d[CGCATATATGCG] (dd1) and d[CGCAAATTTGCG] (dd2) were purchased from Integrated DNA Technologies (Coralville, IA). Stock concentrations were calculated by using the extinction coefficients provided by the supplier. Calf thymus DNA was purchased from Sigma-Aldrich (Milwaukee, WI) and an A_{260} of 12824 M^{-1} cm^{-1} was used to calculate the stock concentration [28]. All samples were prepared in 25 mM potassium phosphate buffer (pH 7.0) and 70 mM KCl, unless otherwise stated. Samples were annealed by heating to 95 ^{o}C in a water bath and then cooling overnight to room temperature in a styrofoam box prior to each experiment.

5.2.2. Synthesis and Characterization of Azacyanines

Aza3 were synthesized as described by Huang et al. [95, 96] and characterized by 1D (^{1}H, ^{13}C) and 2D (HSQC, HMBC) NMR spectroscopy, accurate mass spectrometry, and elemental analysis. Aza3 was also characterized by X-Ray crystallography (crystallization was effected from a saturated methanol-water solution at RT). Extinction coefficients were determined by UV-Vis spectrophotometry, using at least 5 different samples with different concentrations of the compound.

7H-1,13-Dimethyldibenzoimidazolo[1,2-a:2',1'-d][1,3,5]-triazin-6-ium (aza3). ^{1}H NMR (500 MHz, DMSO-d6): δ3.84, (s, 6H), 6.56 (s, 2H), 7.52 (m, 4H), 7.64 (m, 2H), 7.77 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d6): δ29.58, 57.69, 110.89, 111.87, 125.08, 125.30, 128.19, 131.70, 149.68. HR-ESI: Mass observed 290.14066, mass calculated

290.14057 (-0.3 ppm). Elemental Analysis: $C_{17}H_{15.96}N_{4..96}I_1$ (calculated $C_{17}H_{16}N_5I_1$). Extinction coefficient in DMSO: ε_{343} = 47700 $M^{-1}cm^{-1}$. Yield 53%. The crystal coordinates are deposited to Cambridge Crystallographic Data Centre with deposition number 671506.

7H-Dibenzothiazolo[1,2-a:2',1'-d][1,3,5]triazin-6-ium (aza4). ^{1}H NMR (500 MHz, DMSO-d6): δ6.68 (s, 2H), 7.68 (s, 2H), 7.86 (s, 2H), 7.97 (s, 2H), 8.28 (s, 2H). ^{13}C NMR (125 MHz, DMSO-d6): δ61.40, 114.11, 124.83, 125.29, 127.65, 129.58, 139.97, 167.64. HR-ESI: Mass observed 296.03176, mass calculated 296.03162 (-0.5 ppm). Elemental Analysis: $C_{15}H_{10.66}N_{3.00}S_{2.02}I_{0.99}$ (calculated $C_{15}H_{10}N_3S_2.I_1$). Extinction coefficient in DMSO ε_{387} = 32100 $M^{-1}cm^{-1}$. Yield 10%.

7H-3,11-Dimethoxydibenzothiazolo[1,2-a:2',1'-d][1,3,5]-triazin-6-ium (aza5). ^{1}H NMR (500 MHz, DMSO-d6): δ3.90 (s, 6H), 6.58 (s, 2H), 7.86 (dd, 2H), 7.46 (dd, 2H), 7.84 (dd, 2H), 7.88(dd, 2H). ^{13}C NMR (125 MHz, DMSO-d6): δ56.97, 61.40, 109.20, 114.90, 117.20, 126.30, 130.83, 158.92, 165.97. HR-ESI: Mass observed 356.05362, mass calculated 356.05340 (-0.2 ppm). Elemental Analysis: $C_{17}H_{13.91}N_{3.03}S_{2.04}O_{1.99}I_{0.99}$ (calculated $C_{17}H_{14}N_3O_2S_2.I_1$). Extinction coefficient in DMSO: ε_{407} = 32000 $M^{-1}cm^{-1}$. Yield 23%.

5.2.3. Circular Dichroism (CD) and UV-Vis Spectroscopy

CD spectra were acquired on a JASCO J-810 CD spectropolarimeter equipped with a Peltier temperature control unit. Spectra were acquired using a 5 mm strain-free rectangular cell. CD melting profiles were acquired as full spectra from 5 to 95 ^{o}C in 1 ^{o}C steps, at <1 ^{o}C/minute. The buffer spectrum was collected under these conditions and used to correct the sample spectra. Melting curves were obtained by fitting each spectrum

to a two-state model of a linear superposition of the 5 ^{o}C and 95 ^{o}C spectra. UV-Vis absorbance measurements were performed using a HP 8453 UV-Vis diode array spectrophotometer equipped with an Agilent 89090A Peltier temperature control unit. The melting profiles were acquired from 5 to 95 ^{o}C with a rate of <1 ^{o}C/min for all four traces (two heating and two cooling). UV melting profiles were obtained by monitoring the absorbance change at 295 nm, as well as by the aforementioned two-state model.

5.2.4. Fluorescence Titrations

All fluorescence measurements were performed on a Shimadzu RF-5301PC spectrofluorophotometer (Columbia, MD) at 25 ^{o}C. Titrations were carried out by making incremental additions of nucleic acid-1 µM aza3 stock to a solution containing only 1 µM aza3 in buffer. Aza3 was excited at 324 nm and emission spectra were collected from 325 to 700 nm. Excitation and emission slits were both set to 1.5 nm. Equilibrium constants were obtained by fitting the integrated area between 325-600 nm to the least squares equations as described previously [126].

5.2.5. Isothermal Titration Calorimetry

Isothermal titration calorimetry experiments were performed using a Microcal VP-ITC. 5 µl of 200 µM aza3 was injected into 1.4 ml of 100 µM tel24 in the cell. Control titrations were conducted by injecting aza3 into the sample cell containing only buffer under the same conditions. The average heat of dilution was subtracted from the uncorrected heats, and the enthalpy of binding was calculated from the integrated heats and averaged. The experiments were done in duplicate, and the average of the first 50 injections for each run was used to obtain ΔH distribution.

5.2.6. NMR Spectroscopy

NMR experiments were performed on a Bruker DRX-500 at 25 ^{o}C. ^{1}H spectra of the DNA samples collected in H_2O were used to verify proper folding of the G-quadruplex based upon imino proton chemical shifts. The 3-9-19 WATERGATE pulse sequence was used for water suppression. Aza3 titrations were performed by the dissolution of lyophilized aliquots of aza3 into a 450 µl of a 2 mM tel24 sample in D_2O (0.05 equivalents at each step). 128 transients were collected for each 1D ^{1}H. 2D NOESY spectra were collected at selected aza3 to tel24 ratios with 1024 points in F_2 and 512 blocks in F_1. The line broadening of the 1D ^{1}H spectrum was analyzed by fitting resolved resonances by Lorentzian functions with a linear offset. 2D NOESY cross peaks were assigned based on the assignments reported by Luu et al. The software packages SPARKY [127] and SpinWorks [128] were used to analyze 2D spectra. 1D rows were extracted from 2D NOESY data and analyzed by Lorentzian fits. The height changes of tel24 cross peaks in spectra acquired during aza3 titrations were normalized with respect spectra acquired in the absence of aza3.

5.2.7. Surface Plasmon Resonance Measurements

Surface plasmon resonance experiments were performed with a BIACORE 2000 instrument. Biosensor experiments were conducted in degassed HEPES buffer (10 mM HEPES, 100 mM KCl, 3 mM EDTA, 0.00005 v/v of 10% P20 BIACORE surfactant, pH 7.4) at 25 °C. The 5'-biotin-labeled tel24, tel26, d(CGAATTCGTTTTCGAATTCG) (dd3), d(CGCGCGCGTTTTCGCGCGCG) (dd4) and d(CCATATATATATATATAGCCCCCGCTATATATATATATATGG) (dd5) sequences were purchased from Midland Certified Reagent Company with HPLC

purification. The DNA sequences were immobilized on a streptavidin-derivatized gold chip (SA chip from BIACORE) by manual injection of DNA stock solution with a flow rate of 1 µl/min until the SA chip surface was immobilized with the desired amount of DNA. Typically, a series of different ligand concentrations ranging from 10 nM to 10 µM were injected onto the chip with a flow rate of 50 µl/min for a period of 5 min followed by a dissociation period of 10 min. After each cycle, the chip surface was regenerated with a 20-sec injection of pH 2.0 glycine solution and multiple 1-min buffer injections. The number of binding sites and association constants were obtained from fitting plots of RU_{obs} versus C_{free}. To obtain the association constants the data generated were fitted to a three-site interaction model using Kaleidagraph for nonlinear least squares optimization of the binding parameters using the following equation;

$$RU_{obs} = (RU_{max} \times (K_1 \times C_{free} + 2K_1 \times K_2 C^2_{free} + 3K_1 \times K_2 \times K_3 C^3_{free})) / (1 + K_1 \times C_{free} + K_1 \times K_2 C^2_{free} + K_1 \times K_2 \times K_3 C^3_{free})$$

Where K1, K2 and K3 are macroscopic equilibrium constants for three types of binding sites, RU_{obs} is the SPR response at the steady state level, RU_{max} is the maximum SPR response for binding one molecule per binding site, and C_{free} is the compound concentration in solution.

5.3. RESULTS AND DISCUSSION

5.3.1 Binding of Aza3 to Human Telomeric DNA

We investigated the binding of azacyanines to a G-quadruplex sequence, based on the human telomeric repeat d(TTAGGG)n, for which a solution-state structure has been reported: d(TTGGG(TTAGGG)$_3$A) (tel24) [129]. To ascertain discrimination against

duplex DNA, we also conducted binding studies with calf thymus DNA and two isomeric oligonucleotides d(GCGCATATATGCGC) (dd1) and d(GCGCAAATTTGCGC) (dd2). dd2 was specifically designed with an A-tract sequence element that is associated with a narrow minor groove, which typically favors interaction with groove-binding ligands [130].

The binding of aza3 to tel24 was first examined by UV-Vis and CD spectroscopy. In the presence of tel24, the longest wavelength absorption band of aza3 was red-shifted and hypochromic (Figure 5.3) and a weak induced band was observed in the CD spectrum (Figure 5.4). Monitoring the UV-Vis absorption of the G-quadruplex as a function of temperature also revealed that the presence of 1 equivalent of aza3 enhanced the thermal stability of tel24 relative to the unliganded species, raising the T_m from 65 to 67 °C. The enhanced thermal stability of tel24 in the presence of aza3 and the spectral changes observed for aza3 in the presence of tel24 are all characteristic of ligand binding in the chiral environment of DNA.

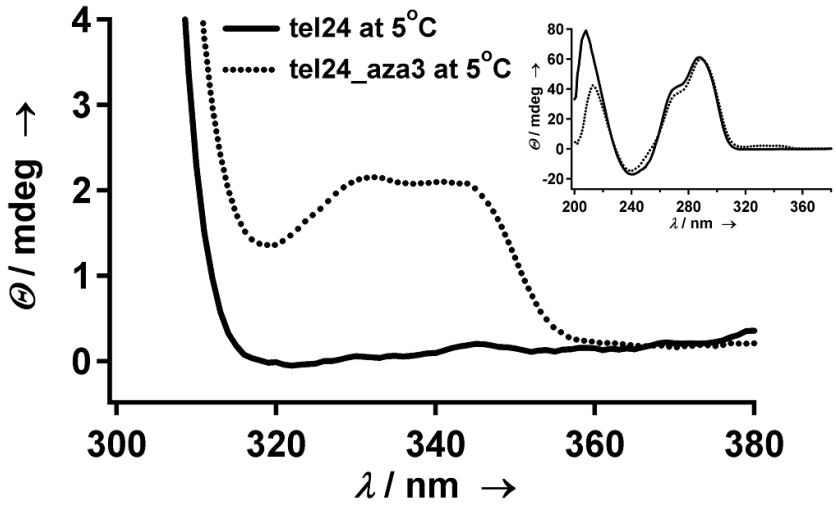

Figure 5.3: Longer wavelength region of the UV-Vis spectra of 1:1 mixture of aza3 and tel24 at 5 ºC (dashed line) and 95ºC (solid line). Samples are 25 µM in strand in 25 mM K-PO_4 buffer with 70 mM KCl at pH 7. Sample with the small molecule has 25 µM aza3.

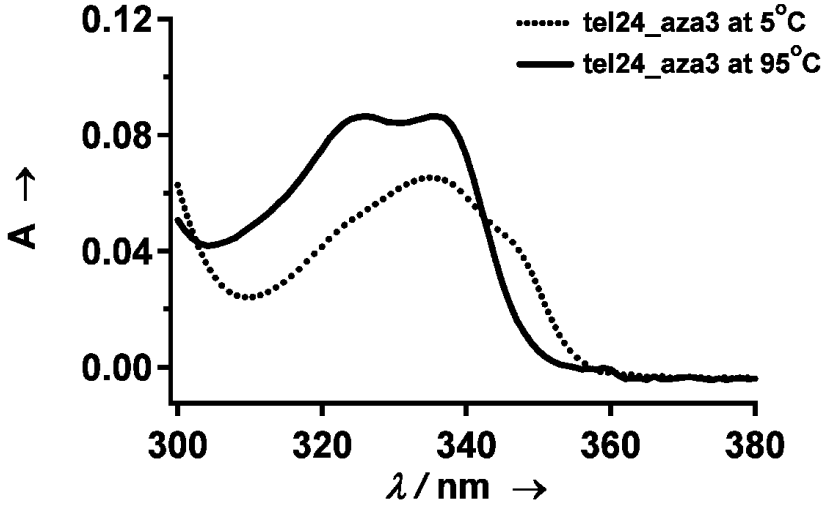

Figure 5.4: Longer wavelength region of the CD spectra of 1:1 mixture of aza3 and tel24 (dashed line) and of tel24 in the absence of aza3 (solid line). Insert shows full CD spectra. Samples are 2.5 µM in strand in 25 mM K-PO_4 buffer with 70 mM KCl at pH 7. Sample with the small molecule has 2.5 µM aza3.

The binding and structural selectivity of aza3 was characterized quantitatively by fluorescence titrations of aza3 with various nucleic acids. The ligand exhibits a profound structural preference, with ca. 100-fold selectivity for G-quadruplex vs. duplex DNA (Table 5.1). Representative binding curves of aza3 to tel24 and CT DNA are given in Figure 5.5 and 5.6 respectively. Aza3 is near what is considered the low end of binding constants (10^3-10^4) for Watson-Crick ligands and near the high end for quadruplex ligands (10^6-10^7); this result supports our size-selection hypothesis. Additionally, aza3 binds more weakly to dd2 than dd1, providing some evidence against groove-binding.

Table 5.1 Association constants determined by fluorescence spectroscopy for aza3 with various DNA samples

Sequence	**Association Constant (M^{-1})**
tel24 d(TTGGG(TTAGGG)$_3$A)	$1.28 \times 10^6 \pm (2.5 \times 10^5)$
Calf Thymus DNA	$9.2 \times 10^3 \pm (1.7 \times 10^3)$
dd1 (d(GCGCATATATGCGC))	$6.13 \times 10^3 \pm (1.2 \times 10^3)$
dd2 (d(GCGCAAATTTGCGC))	$1.14 \times 10^4 \pm (2.0 \times 10^3)$

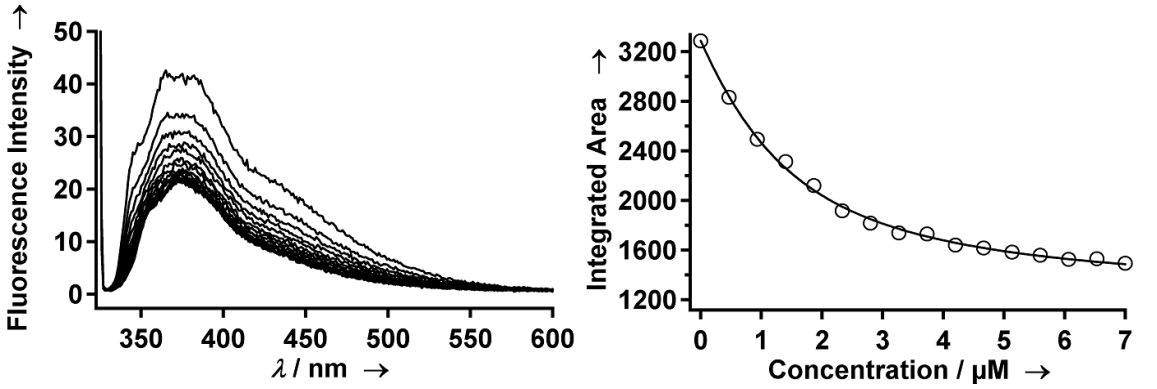

Figure 5.5: a) Fluorescence intensity measurements for 1 μM solutions of aza3 in the presence of increasing concentrations of tel24 at 25 ^{o}C. b) A plot of integrated fluorescence intensities of spectra shown in panel a as a function of DNA concentration (in units of oligonucleotide strand). Curve shown represents a least squares fit to the experimental data (circles). The samples were in 25 mM K-PO_4 buffer with 70 mM KCl at pH 7.

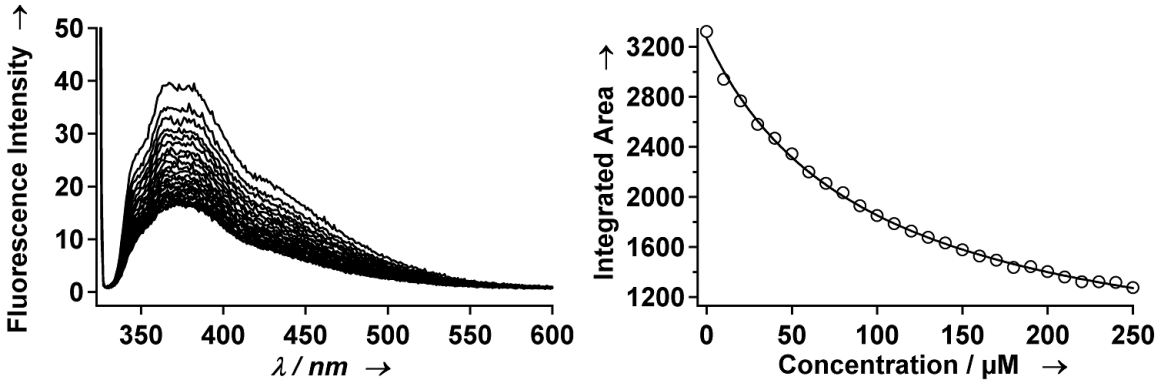

Figure 5.6: a) Fluorescence intensity measurements for 1 μM solutions of aza3 in the presence of increasing concentrations of Calf thymus DNA at 25 ^{o}C. b) A plot of integrated fluorescence intensities of spectra shown in panel a as a function of DNA concentration (in units of oligonucleotide base pair). Curve shown represents a least squares fit to the experimental data (circles). The samples were in 25 mM K-PO_4 buffer with 70 mM KCl at pH 7.

5.3.2 Binding Site Determination

Titration of aza3 into a 2 mM sample of tel24 revealed that the ligand was in intermediate exchange on the NMR timescale. A number of resonances broaden in a site-specific fashion, consistent with the local chemical environment of various residues being perturbed differentially by ligand binding. 1D ^{1}H NMR spectra of tel24 in the presence of increasing concentrations of aza3 are shown in Figure 5.7.

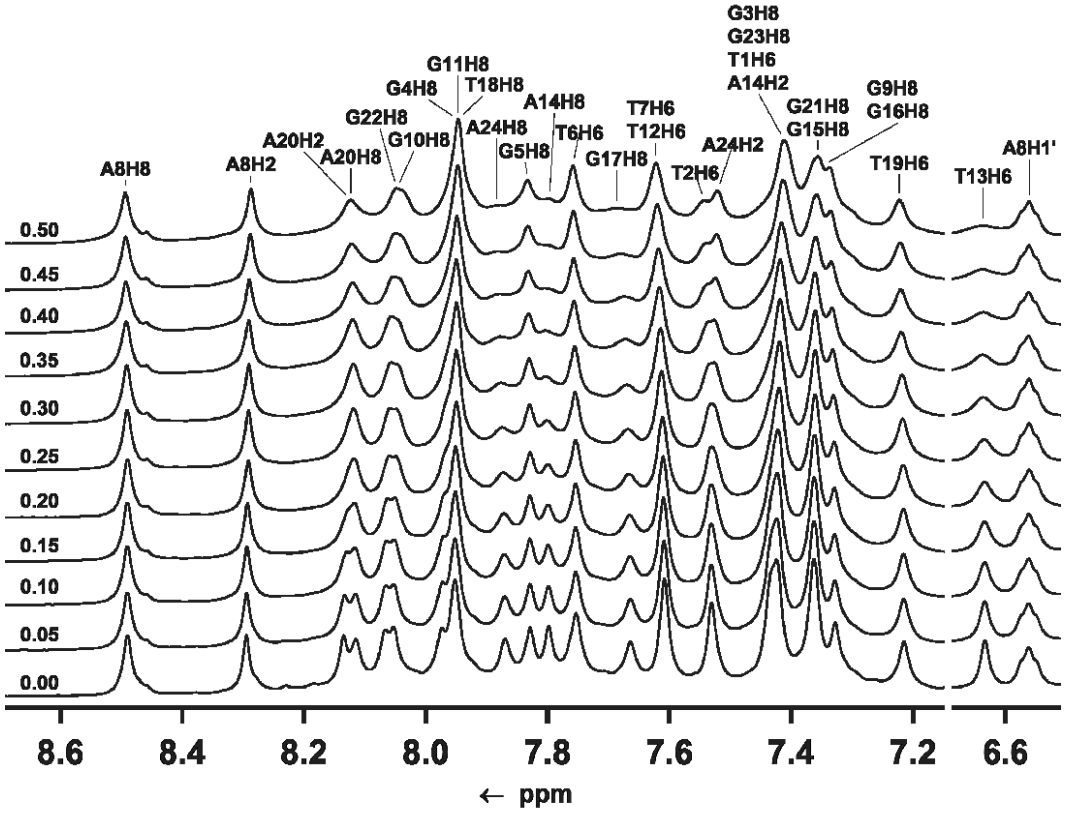

Figure 5.7: Aromatic region of ^{1}H-NMR spectra of 2 mM in strand tel24 in the presence of 0 to 0.50 molar equivalents of aza3. The samples were prepared in 25 mM K-PO_4 buffer with 70 mM KCl and spectra taken at 298 K.

Resonance line broadening was characterized by the fitting of Lorentzian functions to resolved resonances in 1D spectra, and by measuring the heights of aromatic-H1' crosspeaks in 2D NOESY spectra (Figure 5.8). As a qualitative method of determining the ligand binding site, the rate of decrease of the NOESY aromatic-H1' crosspeaks intensities as a function of aza3 concentration was examined (Figure 5.9). The residues were separated into three groups based upon their apparent sensitivity to the presence of aza3. The solution structure of tel24 determined by Luu et. al. [129] is shown in Figure 5.9 with residues colored according to these three groups. This representation clearly illustrates the localized binding of aza3 to tel24.

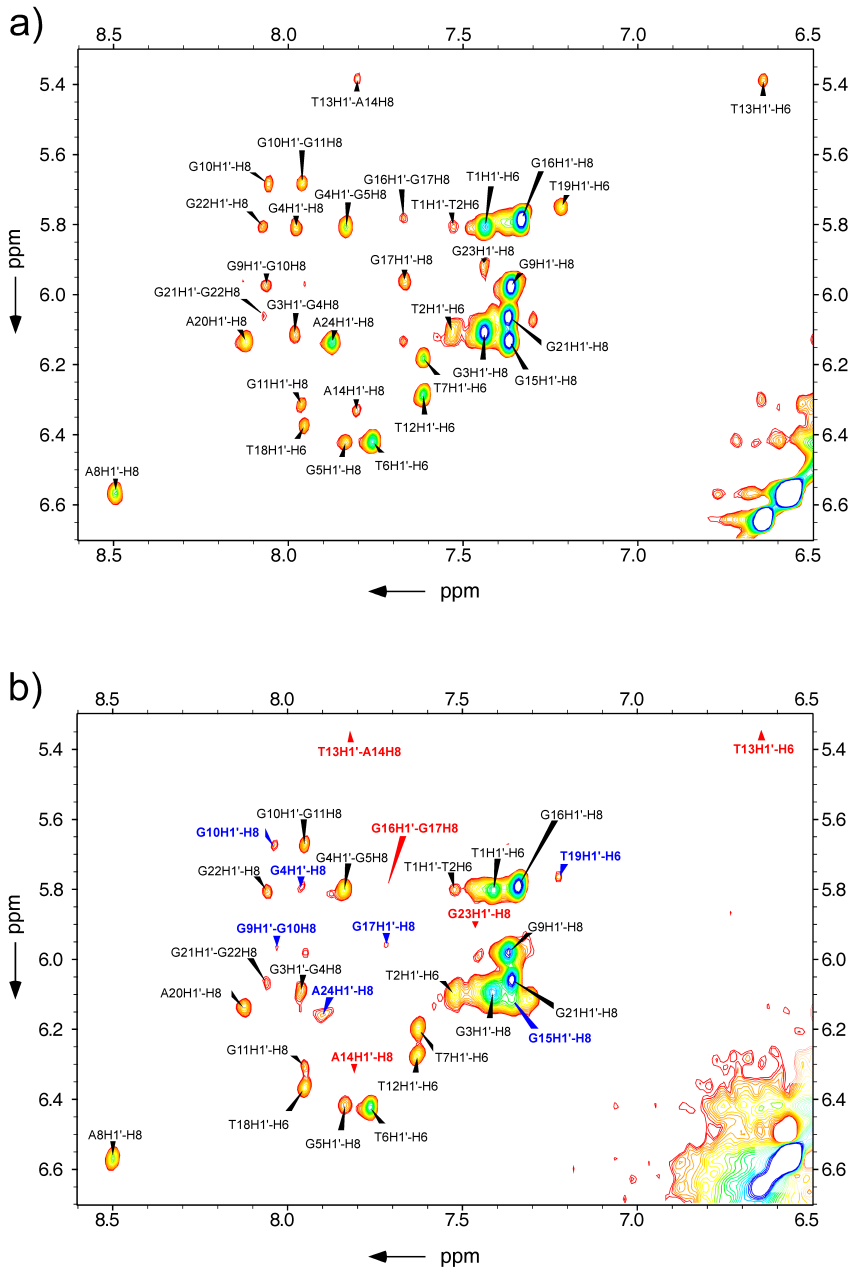

Figure 5.8: a) H6/H8-H1' region of 2D NOESY spectrum of tel24. b) Same region of 2D NOESY spectrum of tel24 in the presence of 0.5 equivalents of aza3. Red arrows indicate cross peaks that disappeared on addition of aza3. Blue arrows indicate cross peaks that exhibited substantial attenuation on addition of aza3. The sample which was 2mM in strand was prepared in 25mM K-PO_4 buffer with 70mM KCl and spectra taken at 298K in H_2O.

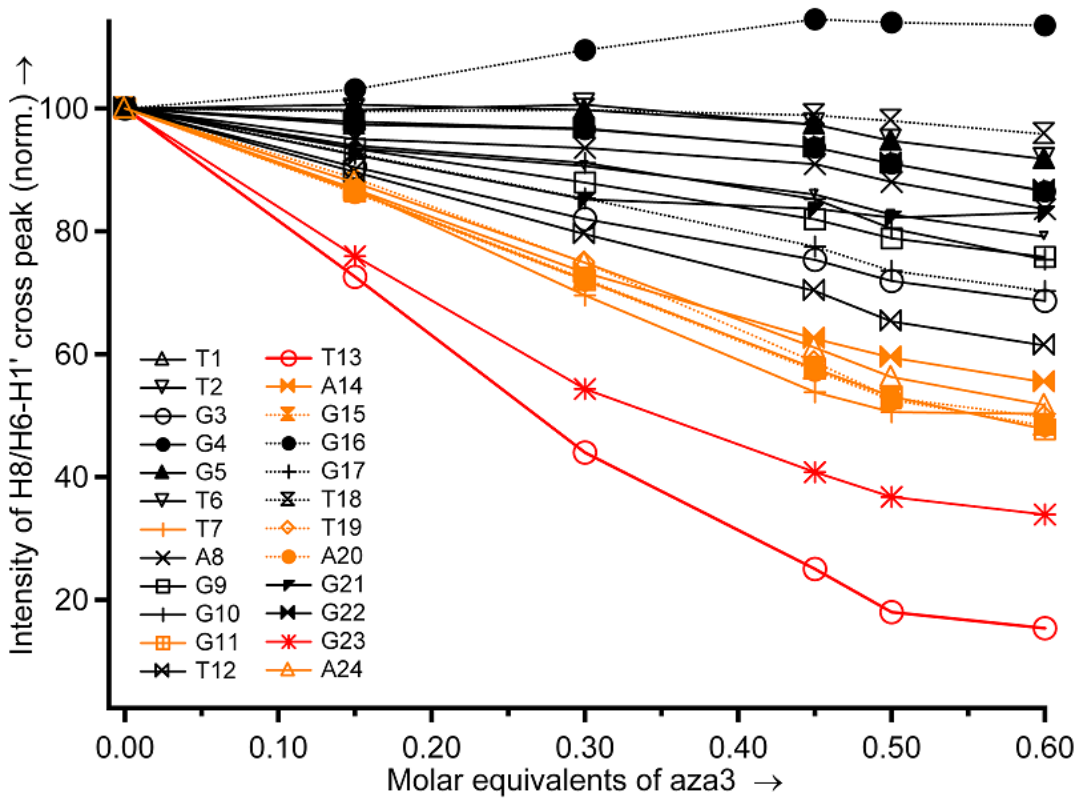

Figure 5.9: Normalized intensities of H6/H8-H1' cross peaks measured in 2D NOESY spectra for tel24 in the presence of various molar equivalents of aza3. All cross peaks measured were in the direct dimension for H1'. The sample which was 2 mM in strand was prepared in 25 mM K-PO_4 buffer with 70 mM KCl and spectra taken at 298 K in H_2O.

Luu et al [129]. reported that an A·T pair is formed by T1 and A20, and a reverse A·T pair is formed by T13 and A24, each being stacked on G-tetrads on opposite sides of the quadruplex. Interestingly, broadening occurs for both the aromatic residues of the external G-tetrads and these A·T base pairs (Figure 5.10), providing support for a mixed intercalation-exterior stacking mode of binding (i.e., between the exterior tetrads and the capping A·T base pairs of the loops). The NMR data indicate the primary binding site lies between the G-tetrad and T13·A24 base pair (Figure 5.10). A secondary weak binding

site is also observed on the opposite site of the quadruplex between the G-tetrad and the T1·A20 pair. The comparison of the peak intensities in 2D NOESY spectra for tel24 and tel24 with 0.5 equivalent aza3 are given in table 5.2 and 5.3. The NMR data do not support a groove-binding mode of interaction, consistent with our results from fluorescence titration of aza3 with dd1 and dd2.

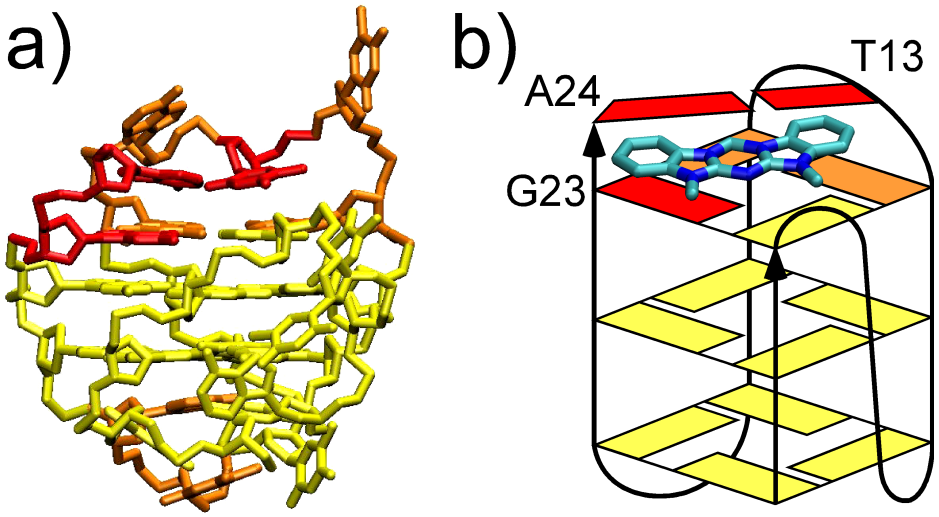

Figure 5.10: a) The solution state structure of Tel24 (PDB ID 2GKU) with residues colored by rate of NOE attenuation per equivalent aza3 (red, high; orange, medium; yellow, low). For clarity, hydrogens and phosphate oxygens are not shown. b) A schematic representation of aza3 bound tel24, based upon NOE attenuation data, in which aza3 is positioned between the A•T base pair of T13 and A24 and G23 of the terminal G-quartet. For clarity, the loop bases are not indicated, except for those of T13 and A24

Table 5.2[a]: Peaks absent in 2D NOESY spectrum of tel24 with 0.5 equivalents aza3 compared to 2D NOESY spectrum of unliganded tel24.

T1 H6/T1 H3'	T1 H3'/T2 H5'*			
T2 H6/ T2 H3'	T2 H1'/ T2 H4'			
T13 H6/ T13 H1'	T13 H6/T13 H2'	T13 H6/T13 H2"	T13 H6/T13 H3'	T13 H6/T13 H4'
T13 H6/T13 H5'*	T13 H6/T13 H5'*	T13 H1'/T13 H2'	T13 H1'/T13 H3'	T13 H1'/T13 H4'
T13 H1'/T13 H5'*	T13 H1'/ A14 H8	T13 H2'/ A14 H8	T13 H1'/A14 H5'*	
A14 H8/A14 H1'	A14 H2'/A14 H3'			
G16 H1'/G17H8	G16 H3'/G17H8			
G17 H8/G17 H1'				
T18 H2"/T19 H6	T19 H2'/A20 H8	T19 H2"/A20 H8		
G23 H6/ G23 H1'	G23 H1'/ G23 H3'			

[a] 5'* represents the non- stereospecifically assigned 5' and 5'' residues.

Table 5.3[a]: Peaks attenuated in 2D NOESY spectrum of tel24 with 0.5 equivalents aza3 compared to 2D NOESY spectrum of unliganded tel24.

T1 H6/T1 H5'*	T1 H1'/T1 H2'			
T2 H1'/T2 H2'	T2 H2'/T2 H3'			
G5 H4'/T6 H6	T6 H6/T6 H3'			
T13 H1'/T13 H2"	T13 H2'/T13 H3'	T13 H2"/T13 H3'	T13 H2"/A14 H8	T13 H3'/T13 H4'
T13 H3'/T13 H5'*	T13 H3'/T13 H5'*			
A14 H8/A14 H2'	A14 H8/A14 H2"			
G15 H8/G15 H1'	G15 H8/G15 H3'			
T19 H6/T19 H1'	T19 H6/T19 H2"	T19 H1'/ T19 H2'	T19 H1'/ T19 H2"	T19 H1'/ T19 H3'
T19 H2'/T19 H3'	T19 H2"/ T19 H3'	T19 H3'/ T19 H5'*	T19 H3'/ T19 H5'*	
G21 H8/ G21 H2'	G21 H3'/G22 H8			
G22 H8/ G22 H3'	G22 H2'/ G23 H8			
G23 H8/ G23 H2'	G23 H8/ G23 H2"	G23 H1'/G23 H2'	G23 H2'/A24 H8	
A24 H8/A24 H1'				

[a] 5'* represents the non- stereospecifically assigned 5' and 5'' residues.

5.3.3 Enthalpy of Binding

"Model-free" ITC titrations (i.e., at very low ligand:binding site ratios) of aza3 into tel24 were performed to better characterize the thermodynamics of the interaction. A ΔH of -2.7±0.14 kcal/mol was obtained (Figure 5.11). With the 1.3×10^6 M^{-1} association constant determined by fluorescence titrations, this datum implies a ΔG of -8.3kcal/mol and a remarkable $-T\Delta S$ of -5.6 kcal/mol. There is a paucity of thermodynamic data on quadruplex-ligand interactions, but such entropy-driven binding has been reported previously for a remarkably different, tetracationic ligand, TmPyP4 [131].

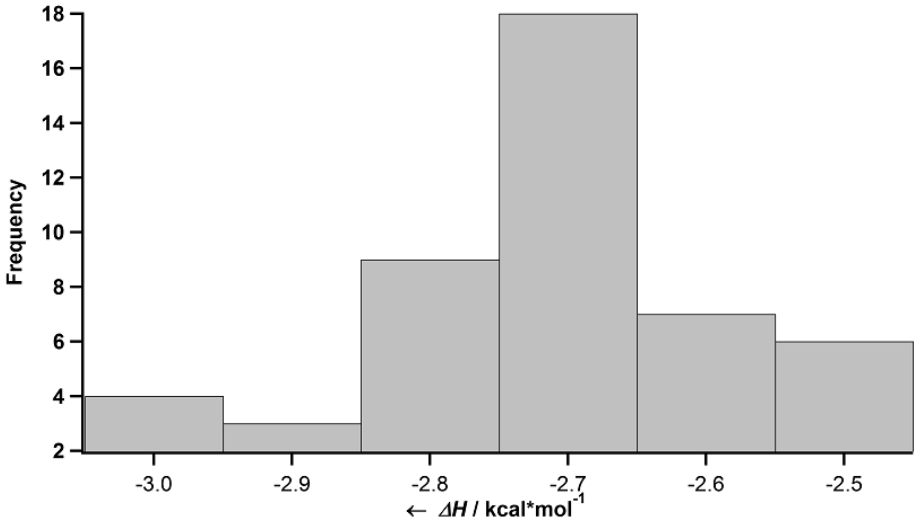

Figure 5.11: The distribution of ΔH values measured by isothermal titration calorimetry for tel24-aza3 binding. The data shown were obtained from the average of two titrations of 50 injections each. 200 µM aza3 was injected into 100 µM in strand tel24 at 22 ^{o}C. The samples were prepared in 25 mM K-PO_4 buffer with 70 mM KCl pH 7.

5.3.4 Broader Spectrum of Binding

Finally, in collaboration with Dr. David Wilson's lab (Georgia State University) the binding of a broader spectrum of azacyanines (aza3, aza4 , aza5) to tel24 was characterized by fitting the steady-state SPR response versus ligand concentration for SPR chips with surface-tethered tel24 DNA. All ligands exhibited tight binding, suggesting that G-quadruplex binding is general to this class of compounds. Association and dissociation kinetics were rapid, as expected from small planar ligands with small substituents. The synthetic accessibility of these compounds suggests substituent variations could be a promising avenue by which to modulate the kinetics for ligand association and dissociation into different ranges for analysis of effects on biological activity.

Additionally, we have tested the binding affinity of all three azacyanine compounds to tel26 (Table 5.4), a twenty-six nucleotide DNA sequence that was also derived from the human telomere sequence and has a similar fold to tel24 [132]. All three azacyanines showed similar strong primary binding affinities to tel26 as well, indicating the observed high affinity is general for this particular quadruplex fold (Table 5.4).

Table 5.4: Association constants determined by SPR for aza3-5 with d(TTGGG(TTAGGG)$_3$A) (tel24), and d(AAAGGG(TTAGGG)$_3$AA) (tel26)

Compound	Sequence	Association Constant (M^{-1})*
Aza3	Tel24	3.1 x10^6
	Tel26	4.1 x10^6
Aza4	Tel24	3.0 x10^6
	Tel26	3.0 x10^6
Aza5	Tel24	2.9 x10^6
	Tel26	4.7 x10^6

*In addition to the association constants shown for the single strong binding site, all compounds have one or two much weaker binding sites (ca. 2 X 105). Association constants for duplex sequences were too low to be determined using the same experimental procedure.

SPR experiments were also performed with intramolecular duplex strands d(CGAATTCGTTTTCGAATTCG) (dd3), d(CGCGCGCGTTTTCGCGCGCG) (dd4), and d(CCATATATATATATATAGCCCCCGCTATATATATATATATGG) (dd5) (hairpin loop regions are underlined). These strands contain a variety of sites known to favor various modes of ligand binding in duplex DNA, including A-tracts and both pyrimidine-purine steps. The azacyanine ligands bind poorly to all duplex DNA strands investigated and the SPR data cannot be fit to an exponential binding curve (Figure 5.12); the upper bound for the *K*a in each aza3 and aza4-dsDNA pair is therefore less than 10^4 M^{-1}, consistent with our fluorescence data. Remarkably, no aza5-duplex interaction was detected under the SPR conditions used (Figure 5.13). Fluorescence binding data conform that these ligands exhibit marked selectivity for quadruplex over dsDNA over 100 fold in the case of aza5.

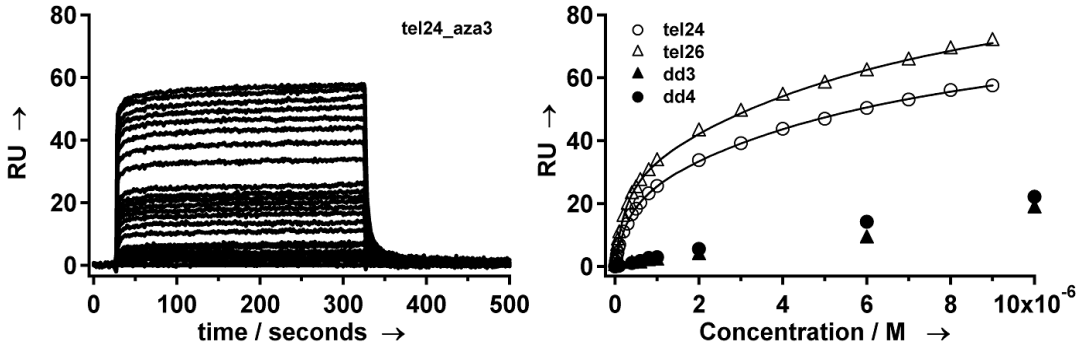

Figure 5.12: a) SPR sensorgram for binding of aza3 to tel24. b) Response averages from the steady state regions of the sensorgrams of aza3 binding to different DNA sequences (open triangles: tel26, open circle: tel24, filled circle:dd4, filled triangle: dd5). The curves represent the best fit of the response averages from the steady state regions of the sensorgrams to three different binding site models. The samples were prepared in HEPES buffer at pH 7.4 and the experiments were conducted at 25 ^{o}C. See materials and methods for further information.

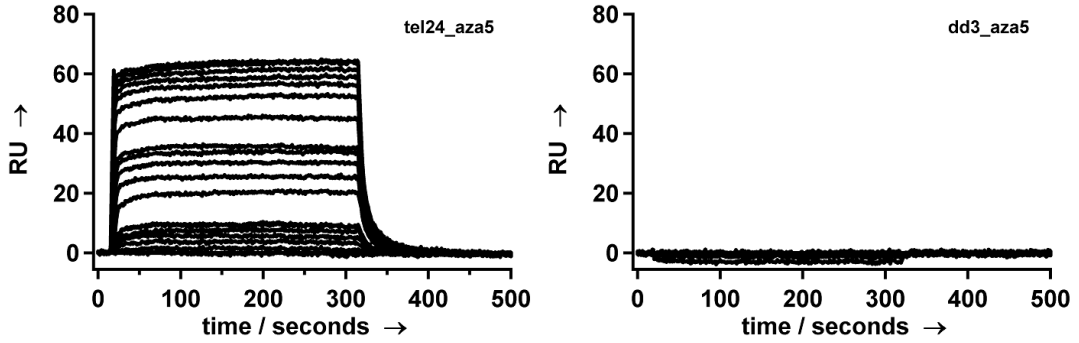

Figure 5.13: SPR sensorgrams for binding of aza5 to a) tel24 and b) dd3. The samples were prepared in HEPES buffer at pH 7.4 and the experiments were conducted at 25 ^{o}C. For more information see materials and methods.

5.3. CONCLUDING REMARKS

G-quadruplex ligands hold great promise for use in the treatment of human disease. A common and valid critique of their medicinal utility relates to the gravity and wide range of off-target effects associated with binding to Watson-Crick DNA. Only by ameliorating such effects is the G-quadruplex a viable drug target. We have shown that a bispurine analogue, which is apparently slightly too large to intercalate a purine-pyrimidine base pair, exhibits strong and selective association with a G-quadruplex. Kurth's synthetic route to azacyanines provides a general and perhaps the most practical means thus far for the preparation of selective G-quadruplex ligands. A promising avenue of future research will be the decorating of azacyanines with the appropriate functional groups to bind specific promoter-related quadruplexes selectively, as Hurley and coworkers have shown is possible with quarfloxin [133, 134].

CHAPTER 6

CONCLUDING REMARKS AND FUTURE DIRECTIONS

6.1. INTRODUCTION

It has been more than 50 years since the discovery of DNA as the carrier of our genetic information. The discovery of the DNA structure has altered the course of science and still the nucleic acids are under massive investigation. Each and everyday we learn a new function, a new structure of nucleic acids in the genetic machinery. We discover a new way of using them to perform new functions. Today, even the definition of "central dogma" of molecular biology has transformed, with the discovery of RNA as being the genetic carrier in RNA viruses and retroviruses [36]. Watson-Crick base pairing will always be heart of the nucleic acids. However, the interest in non-canonical base pairing is also increasing at a significant pace due to their role from biologically relevant RNA structures to the prebiotic world.

Here, complementary to other published results we have demonstrated that small molecules can act as molecular switches, selectively recognize a specific nucleic acid structure and alter the nucleic acid secondary structure under appropriate conditions [29, 31, 37, 59, 65]. They can act as molecular midwifes and assemble nucleic acids into different secondary structures which are otherwise unstable. Of course, the structure of both the small molecule and the base pairing are equally important in this selective molecular recognition.

6.2. PURINE-PURINE BASE PAIRING

The assembly of the homod$(A)_n$ into anti-parallel double helix is only an example of the "molecular midwife" hypothesis. And it begs the question about the other possible purine-purine base pairings, and their intercalation mediated assembly. The importance of the purine-purine base pairing actually lays on the grounds of the poor stacking of pyrimidine bases in aqueous solution [61, 135]. It is worth considering the hypothesis that the chemical evolution began with purine-purine base pairs. As described in Chapter 1, before the formation of the proto-RNA, the nucleic acid bases need to assemble into columnar stacks by using the molecular midwifes as the possible templates. Due to their large size, the stacking interactions of the purine bases with the molecular midwife's will always be favoured compared to the pyrimidine bases. This hypothesis of purine-purine base pairs has been supported by several prominent researchers over the past forty years for several reasons, in part as a solution to the problem of poor pyrimidine base stacking [36, 136-140].

The discovery of the formation of duplexes with the purine-purine base pairs goes back to 1958. Rich et. al. suggested the formation of duplexes by A·I base pairing in mixtures of polyadenylic and polyinosinic acid [141], however, the reinvestigation of these mixtures suggested the triplex structure formation by A·I·I triplets (which contains A·I base pairs) rather than A·I duplex [142]. The widespread investigation of all the possible purine-purine base pairing has been presented by Eschenmoser and colleagues [143]. They have provided evidence for duplex formation by different purine-purine base pairings in hexopyranosyl-(6'→ 4')-oligonucleotide (homo-DNA) systems [143]. They have categorized the purine bases into two classes, adenine like and hypoxanthine like,

where each base in a certain category can form at least two H bonds to the base in the other category (Figure 6.1).

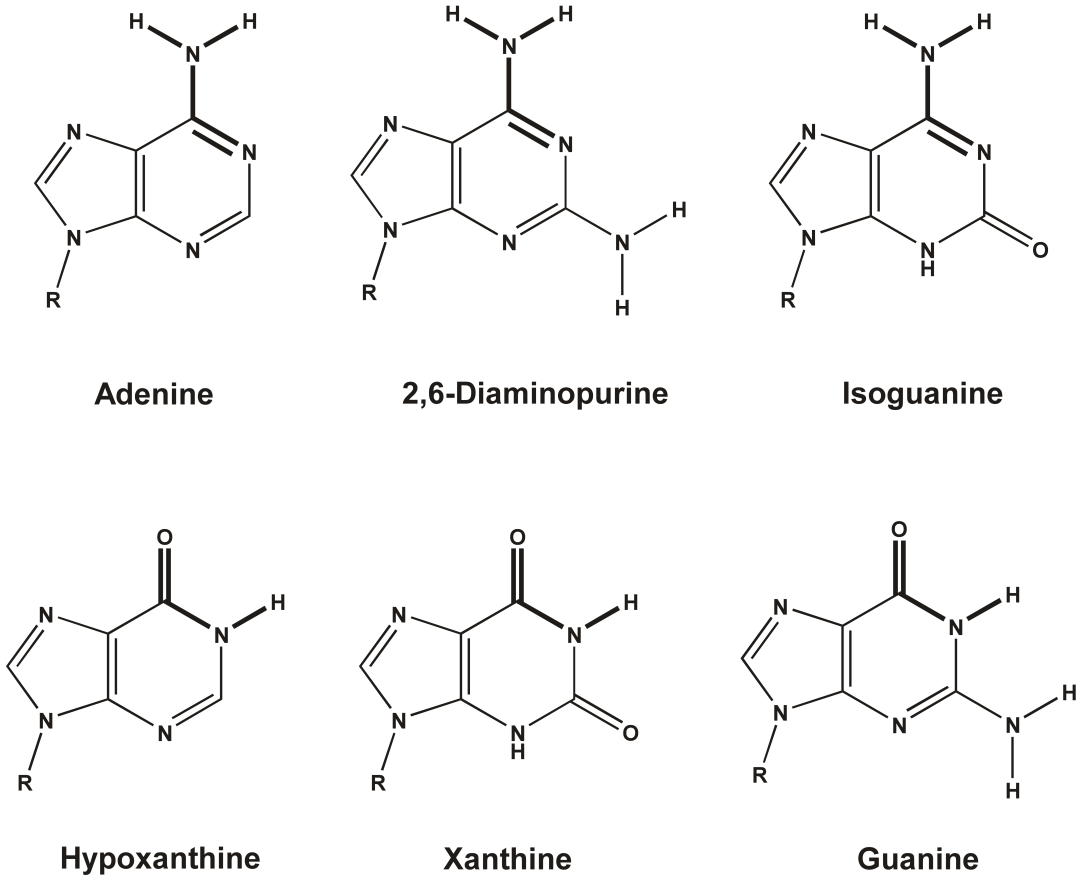

Figure 6.1: Structures of adenine (top panel) and hypoxanthine (bottom panel) like purine bases. Common groups dictating the H-bond formation are highlighted.

By thermal denaturation studies they have reported the formation of very stable duplexes especially containing D·X (diaminopurine·xanthine) and G·isoG (guanine·isoguanine) base pairs. Their results are summarized in table 6.1 and the structures of D·X and G·isoG are given in Figure 62. Most recently, Battersby et al. have

demonstrated duplex formation by DNA oligonucleotides containing a mixture of A·I(adenine·inosine) and G·isoG base pairs [140].

Table 6.1: T_m of the purine-purine hexamers of homo-DNA. Samples are between 2 µM-15 µM in strand concentration in 150 mM NaCl, 10 mM Tris pH 7. Table is adapted from Eschenmoser et. al. [144].

	Adenine	**Inosine**	**Diaminopurine**	**Xanthine**	**Isoguanine**	**Guanine**
Adenine	47	<10	41	14	43	<15
Inosine	<10	–	<5	–	NA	<15
Diaminopurine	41	<5	36	–	39	<15
Xanthine	14	–	–	–	16	18
Isoguanine	43	NA	39	16	42	61
Guanine	<15	<15	<15	18	61	38

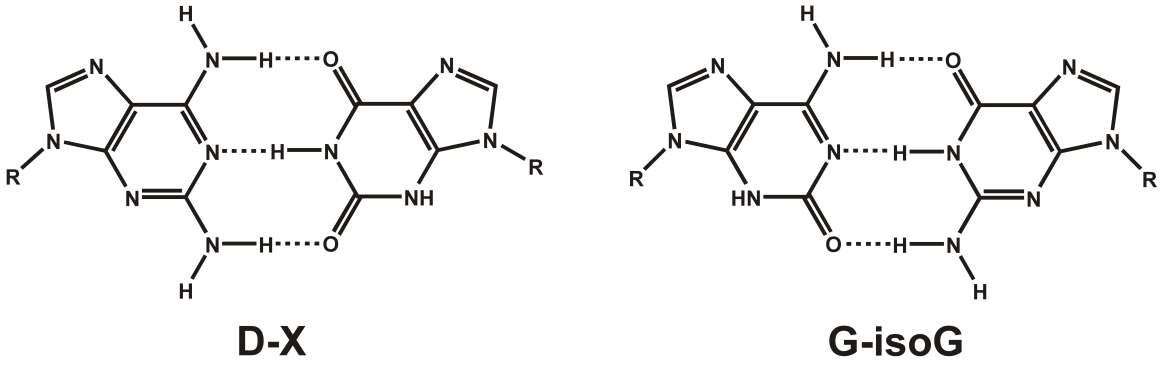

Figure 6.2: Structures of the most stable base pairing systems observed in homo-DNA. D-X and G-isoG refer to diaminopurine-xanhtine and guanine-isoguanine respectively.

Understanding the fundamentals of purine-purine systems and their interactions with small molecules not only will help us to have an understanding of the midwife hypothesis, but will also open a new era in non-canonical systems which might be proven useful in nanotechnologies. The preliminary results on different types of purine-purine

duplexes and their interactions with small molecules are reported in the following sections.

6.3. ADENINE·INOSINE

6.3.1. Materials and Methods

Coralyne chloride is purchased from Sigma and the concentrations of stock solutions in H_2O are determined spectrophotometrically by using $\varepsilon_{420} = 14500\ M^{-1}\ cm^{-1}$. Poly(A) and poly(I) were purchased from Sigma and the $\varepsilon_{258} = 9800\ M^{-1}\ cm^{-1}$ $\varepsilon_{248} = 10000\ M^{-1}\ cm^{-1}$ base is used to calculate the concentrations. 5'-d(IAAIAIIAA)-3' was synthesized on an Expedite 8909 DNA synthesizer. dI-CE Phosphoramidite is purchased from Glen Research and used as received. 5'-d(GGIACC)-3 is purchased from Integrated DNA technologies. The concentration of 5'-d(IAAIAIIAA)-3' and 5'-d(GGIACC)-3' were calculated based on the extinction coefficients supplied by IDT (IDT oligoanalyzer).

Thermal denaturation analysis are performed by using CD spectroscopy from 5-95 ^{o}C, 1 ^{o}C /min. Samples are all in 60 μM in base and 15 μM in small molecule when it is present. Other conditions are given in figure legends.

NMR experiments were performed on a Bruker DRX-500. ^{1}H spectra of the DNA samples collected in H_2O using the 3-9-19 WATERGATE pulse sequence for water suppression.

6.3.2. Binding of Coralyne to Poly(A)·Poly(I)

By using CD spectroscopy we have investigated the interaction of coralyne with polyI, poly(A)·poly(I) and poly(A)·poly(I)$_2$. All of the samples were annealed first overnight, and the coralyne is added just before the denaturation studies. The samples were heated and cooled down from 5-95 ^{o}C with 1 ^{o}C / min steps. Monitoring the changes in CD spectra revealed a T_m of 38 ^{o}C for poly(I)-coralyne, 40 ^{o}C poly(A)·poly(I), 40 ^{o}C poly(A)·poly(I)$_2$. Very broad transitions also observed for (poly(A)·poly(I))-coralyne, (poly(A)·poly(I)$_2$)-coralyne samples between 5-50 ^{o}C. Cooling down the samples to 5 ^{o}C after heat cycling resulted in yellow precipitate-like structure formation in all the coralyne samples. There was a clear change in CD spectra from the original under these conditions (Figure 6.3). These precipitate-like structures were able to float in solution rather than precipitating to the bottom of the tubes. No precipitation is observed in the sample which has only coralyne in it under the same buffer conditions. Also no spectral change is observed in samples without coralyne (Figure 6.3). The samples with coralyne were centrifuged and the supernatant was separated from the precipitates. Thermal denaturation data was collected on the supernatant of these samples again. The CD spectra of all the samples looked alike with a transition around 70 ^{o}C (Figure 6.4).

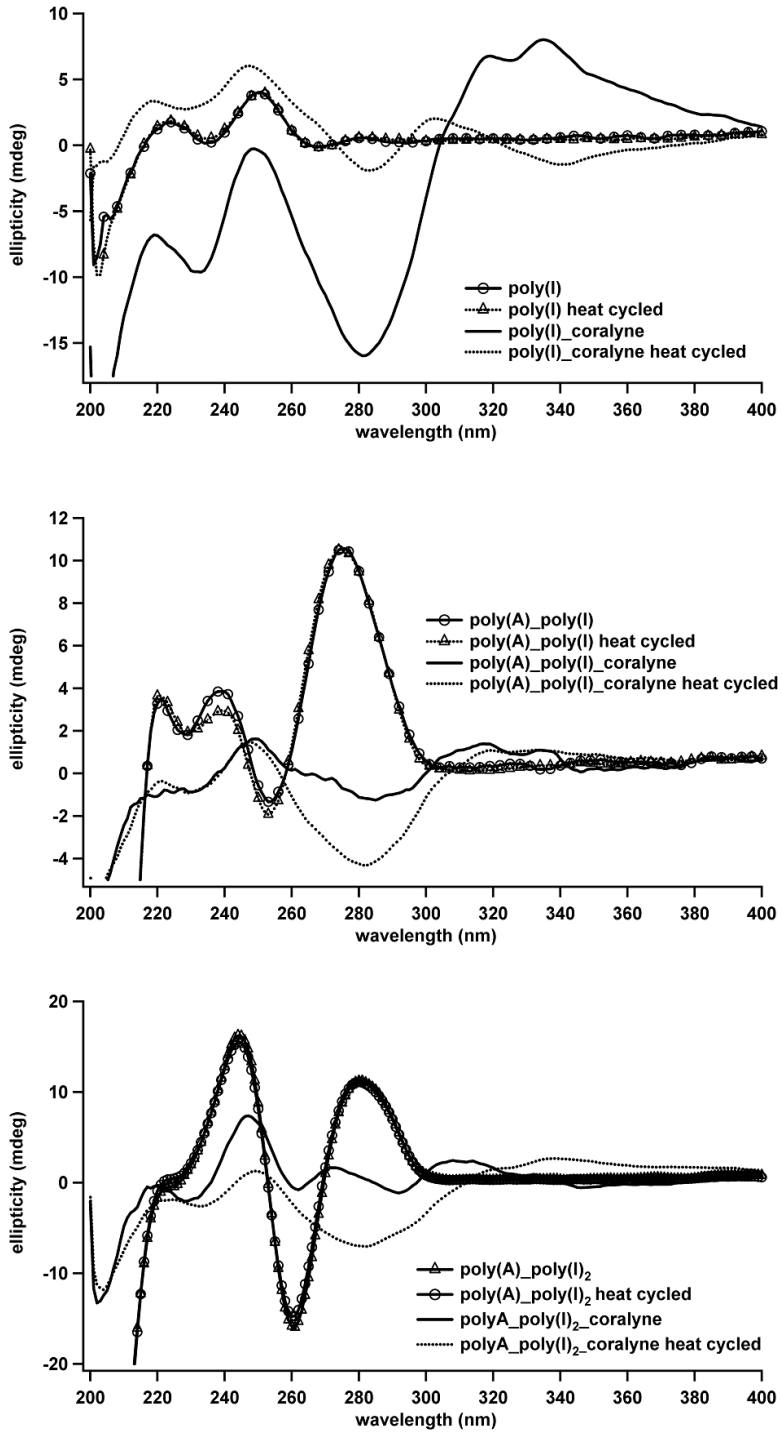

Figure 6.3: CD spectra of poly(I), poly(A)·poly(I) and poly(A)·poly(I)$_2$ in the presence and absence of coralyne. The spectra represented are taken at 5 ^{o}C before and after heat cycling. Samples are 60 µM in base, base pair and base triplet respectively. Coralyne concentration is 30 µM in samples with coralyne. All samples are prepared in 1X BPE at pH 7 with 100 mM NaCl.

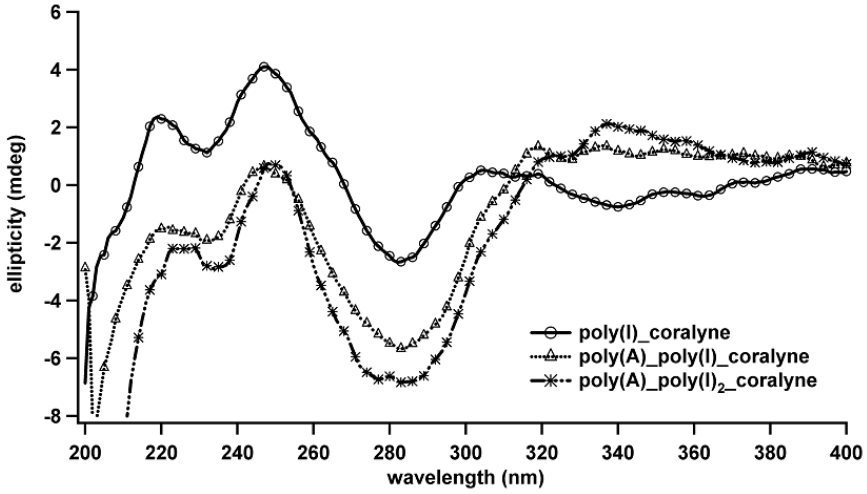

Figure 6.4: CD spectra of the supernatant of poly(I), poly(A)·poly(I) and poly(A)·poly(I)$_2$ samples shown in Figure 6.3 in the presence of coralyne at 5 ^{o}C.

Based on the observations of precipitations in all samples in the presence of the coralyne and the similarities of the CD spectra of the supernatant of poly(I), poly(A)·poly(I) and poly(A)·poly(I)$_2$ samples even though they have different concentrations of poly(I), clearly indicates the interactions of coralyne with poly(I). Coralyne results a precipitate-like structure formation in poly(I) and causing the dissociation of the poly(I) from poly(A) duplex/triplex. Further investigation is needed to be able to understand the cause of this precipitation.

6.3.3. Binding of Coralyne to Adenine-Inosine Oligonucleotides

We decided to further investigate adenine-inosine base pairing and the binding of small molecules to these nucleic acid structures in an oligonucleotide system. We choose 5'-d(IAAIAIIAA)-3' as our model system. As mentioned previously in introduction,

triplex formation is more stable over duplex formation in poly(A), poly(I) mixtures, where adenine as the middle strand is able to form both Watson-Crick and Hoogsteen base pairing with the poly(I) strands. It has been suggested that the formation of the Hoogsteen base pairing is required for any base pairing to occur in this system [142]. Keeping this in mind, we chose a mixed sequence. Since inosine can not form Hoogsteen base pairing when it is in the middle of the strand, this oligonucleotide should prefer to form a duplex rather than a triplex.

Initial thermal denaturation studies in CD resulted in no clear transition both in the presence and absence of the small molecules. The spectra of all the samples looked alike, and no clear induced CD band is observed upon the addition of small molecule (Figure 6.5). The change in CD intensity was quite remarkable and reversible upon denaturation. However, when change in CD intensity is observed at varying temperatures, the T_m curves look like the destacking of the single strands (Figure 6.5). Addition of further salt (either 1M NaCl or 50 mM $MgCl_2$) to stabilize this duplex didn't cause any change in CD spectra.

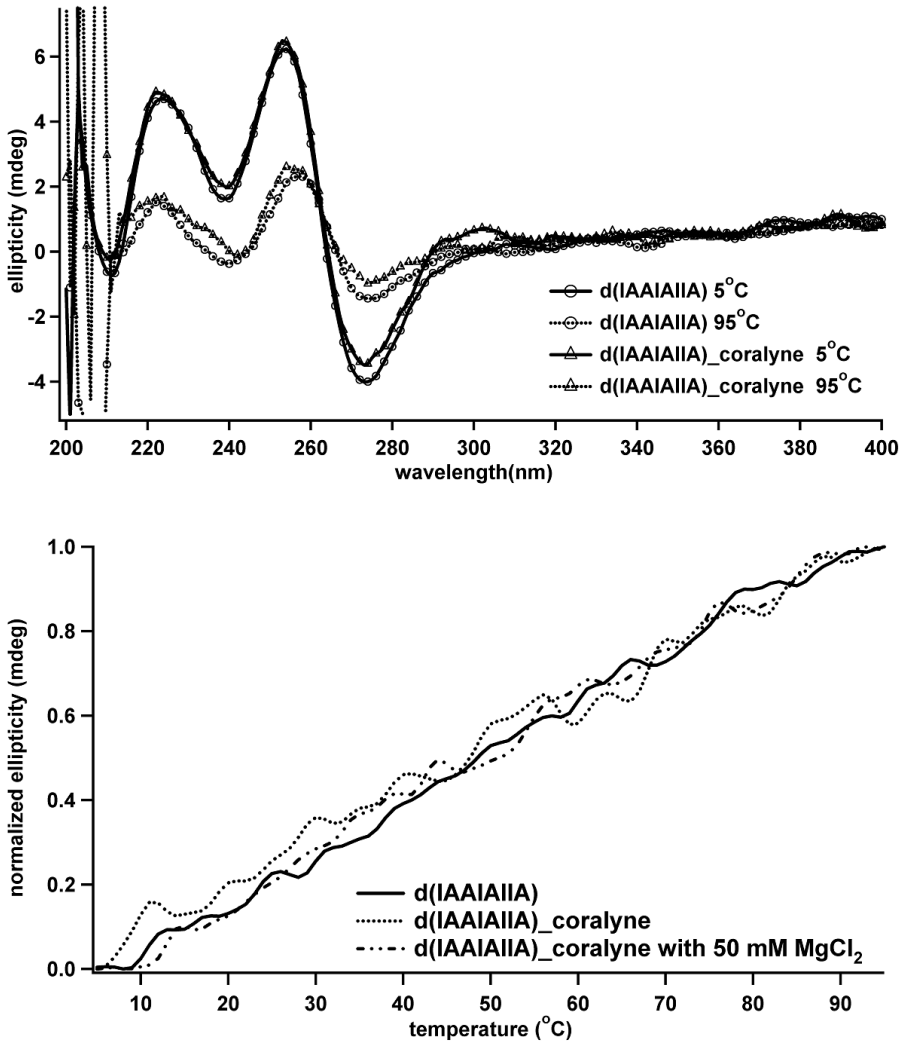

Figure 6.5: CD spectra of d(IAAIAIIA) oligonucleotide in the presence and absence of coralyne at 5 ^{o}C (solid line) and 95 ^{o}C (dashed line) (upper panel). The change in ellipticity at 273 nm with varying temperatures for d(IAAIAIIA), d(IAAIAIIA)-coralyne and d(IAAIAIIA)-coralyne with additional 50 mM $MgCl_2$ (bottom panel). The samples are all in 60 µM in base and prepared in 1XLiCacodylic buffer with 100 mM $LiCl_2$ at pH 7. The samples with coralyne have 15 µM coralyne in them.

On the other hand, the analysis of the d(IAAIAIIA) sample at higher concentrations in NMR reveals the formation of a secondary structure. There were observable very broad imino protons between 276 K and 288 K at pH 5.5. The spectrum of the sample doesn't change by bringing the pH to 7 except in the imino region, most

likely due to the fast exchange. The imino region and the full ^{1}H NMR spectra are given in Figure 6.6.

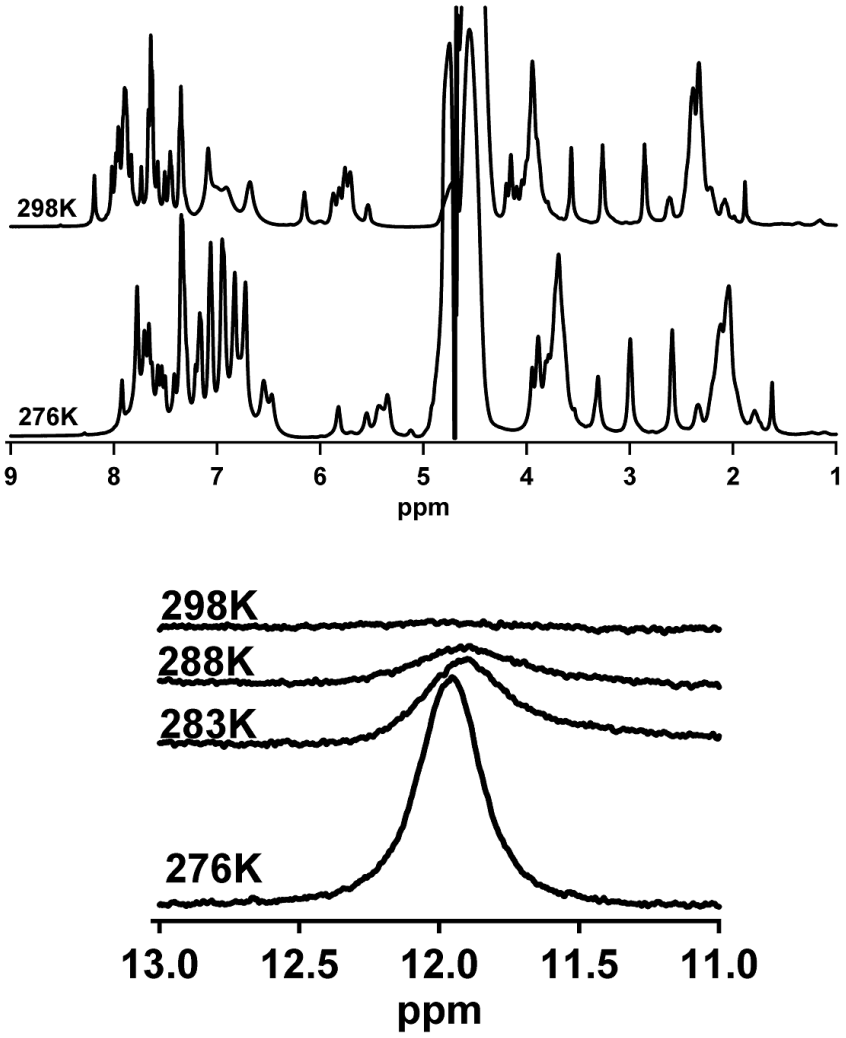

Figure 6.6: ^{1}H NMR spectrum of 6 mM d(IAAIAIIA) (upper panel). Only the imino region of the same sample at various temperatures is shown in bottom panel. The samples are at pH 5.5 with 50 mM NaCl.

Even though no clear transition is observed in CD spectroscopy, our NMR experiments states the otherwise. One obvious reason can be the shift of the equilibrium between the single strands and the duplex with increasing concentration in NMR experiments. Also, one should keep in mind that even as single stranded; these purine oligonucleotides will have very favorable stacking interactions in solution. And the hypochromicity change in spectroscopy might not be as much as a normal duplex going from duplex to single strands. Thus, the destacking of the purine strands might mask the transition of the secondary structure change. Battersby et. al. also observed this effect in their system. Monitoring the melt of purine-purine duplexes with spectroscopy didn't result in any cooperative transition, unless they tagged one of the strands with a fluorophore [145].

6.3.4. Binding of Coralyne to Adenine-Inosine Base Pairs Incorporated in Watson-Crick Duplexes

The formation of a A·I base pairing incorporated into complementary DNA duplexes as mismatches is observed by using NMR spectroscopy and X-ray crystallography [146, 147]. Different types of base pairing arrangements are observed under these conditions. The NMR spectroscopy of 5'-d(GGIACC)-3' revealed a A(anti)·I(anti) base pairing, while crystal structure of 5'-d(CGCIAATTAGCG)-3' revealed A(syn)·I(anti) base pairing between the two. The observed structures are given in Figure 6.7. The analysis of the backbone revealed an altered duplex structure due to the incorporation of two mismatches in 5'-d(GGIACC)-3' duplex. On the contrary, very little deformation is observed in 5'-d(CGCIAATTAGCG)-3' duplex.

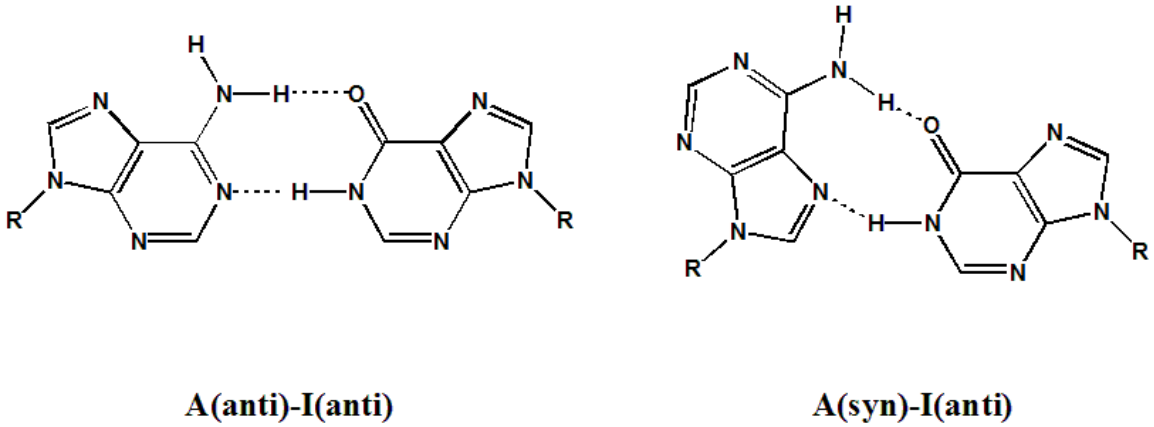

Figure 6.7: Different base pairings observed in A·I by NMR (left panel) and X-ray (right panel) [146, 147]. A·I refers to adenosine and inosine.

We decided to investigate the interactions of 5'-d(GGIACC)-3' with small molecules by NMR spectroscopy. The preference for the sequence was due to having 2 A·I base pairings next to each other, suitable for small molecule intercalation.

Our initial NMR experiments with the revealed the formation of the duplex at 276K. Three major imino protons are observed as expected [147] (Figure 6.8). In addition, watergate experiments also confirmed the duplex structure where, NOEs are observed from imino protons to H6/H8 region of the spectra (Figure 6.8). However, both in 1D and 2D NMR, a second group imino protons observed in the same region.

Studies regarding to identify the structure of 5'-d(GGIACC)-3' duplex and its interactions with small molecules are still going on in our lab.

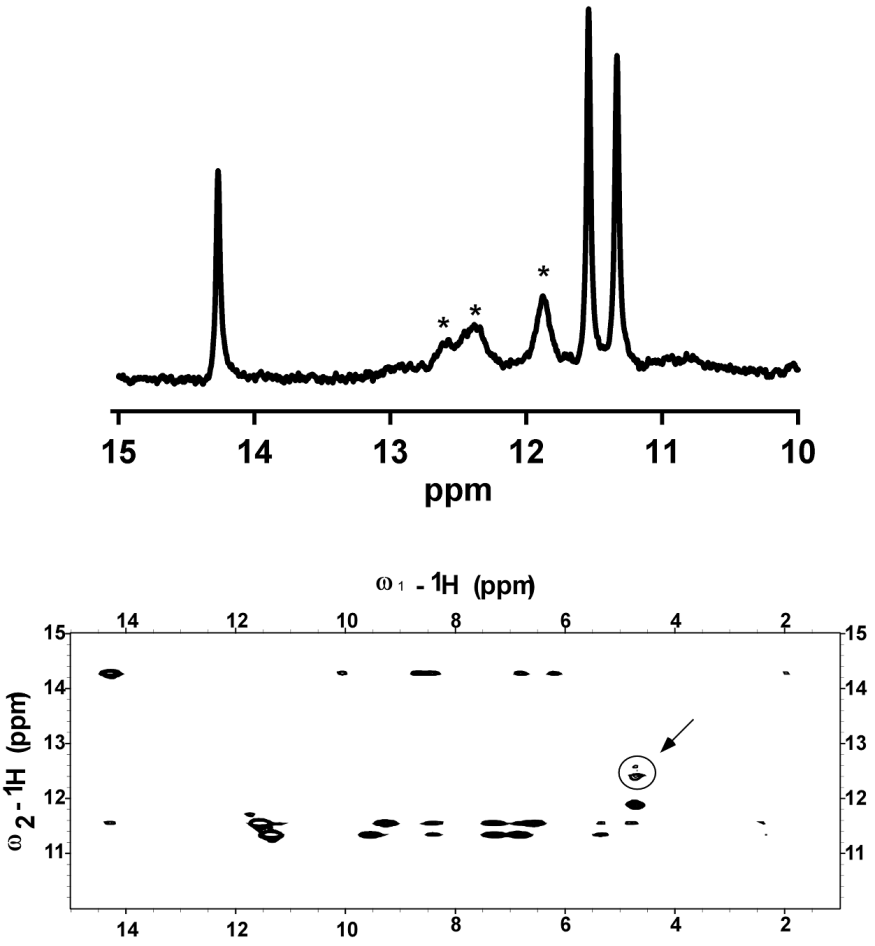

Figure 6.8: Imino region of the ^{1}H-1D NMR (upper panel) and the imino to H6/H8 region of NOESY spectra of 5'-d(GGIACC)-3' duplex. Sample is 2 mM in strand concentration at pH 5.5 [146, 147]. Extra peaks are labeled with asterisk in the ^{1}H-1D and with an arrow in ^{1}H-2D NMR.

6.4. GUANINE·ISOGUANINE

Isoguanine is a non canonical purine base which can be formed spontaneously by the oxidation of adenine in biological systems [148]. It has oxygen at the 2 position of the adenine. It can base pair with isocytosine in order to form base pairing similar to G·C

[149]. It has very similar physical properties to guanine such that it can also form tetraplex as well as pentaplex structures by self-association [150-154]. Eschenmoser has been demonstrated that G·isoG can form very stable structures in homo-DNA hexamers. We have synthesized self-complementary deoxyoligonucleotides to examine this base pairing and its interactions with small molecules.

6.4.1. Materials and Methods

Oligonucleotides were synthesized on an Expedite 8909 DNA synthesizer. dmf-isodG-CE Phosphoramidite is purchased from Glen Research and used as received. All of the oligonucleotides were purified by HPLC, desalted and characterized by mass spectroscopy. The extinctions coefficients of the oligonucleotides are calculated by using IDT (Integrated technologies) oligo analyzer based on the nearest neighbor parameters.

Synthesized Oligonucleotides and the extinction coefficients are as following;

5'-d(isoG-G-isoG-G)-3'; $\varepsilon_{260} = 28000\ M^{-1}\,cm^{-1}$

5'-d(isoG-G-isoG-G-isoG-G)-3'; $\varepsilon_{260} = 41300\ M^{-1}\,cm^{-1}$

5'-d(isoG-G-isoG-G-isoG-G-isoG-G)-3'; $\varepsilon_{260} = 54600\ M^{-1}\,cm^{-1}$

Thermal denaturation analysis are performed by using CD spectroscopy from 5-95 ^{o}C, 1 ^{o}C /min. Samples are all in 60 µM in base in 1X LiCacodylic buffer with 100 mM LiCl, unless otherwise mentioned. The samples with the small molecule have 15 µM of aza3.

6.4.2. Results and Discussion

The CD spectra of all the isoG·G oligonucleotides in the presence and absence of aza3 are given in Figure 6.9. There wasn't any change in the CD spectra upon small molecule binding, no induced CD band formation is observed. Only T_m of d(isoG-G-isoG-G) oligonucleotide could be obtained by monitoring the changes in intensity of the CD bands. The transitions observed in the two other strands were not complete even at 95 ^{o}C under these conditions, 1X LiCacodylic buffer with 100 mM LiCl. This result suggested the possibility of quadruplex formation even in these mixed isoG·G oligonucleotides, similar to what has been observed before [150-154]. We tested this idea by adding KCl to the d(isoG-G-isoG-G) since the role of K^+ in stabilizing nucleic acids has been well known [112]. While adding only 10 mM KCl resulted in an increased cooperativity of the denaturation and a T_m shift of about 25 ^{o}C in the d(isoG-G-isoG-G) tetramer; adding 100 mM KCl shifted the T_m even to higher temperature which was not detectable under these conditions (Figure 6.10). It clear that the isoG·G is able to form very stable quadruplexes, and no interaction between aza3 and this quadruplex structures is observed under these conditions. In order to favor the duplex formation over quadruplex, studies of mixed sequences incorporating isoG·G base pairs with other purine-purine base pairs are under further investigation. The interactions of this base pairing with other small molecules under different conditions will also be explored.

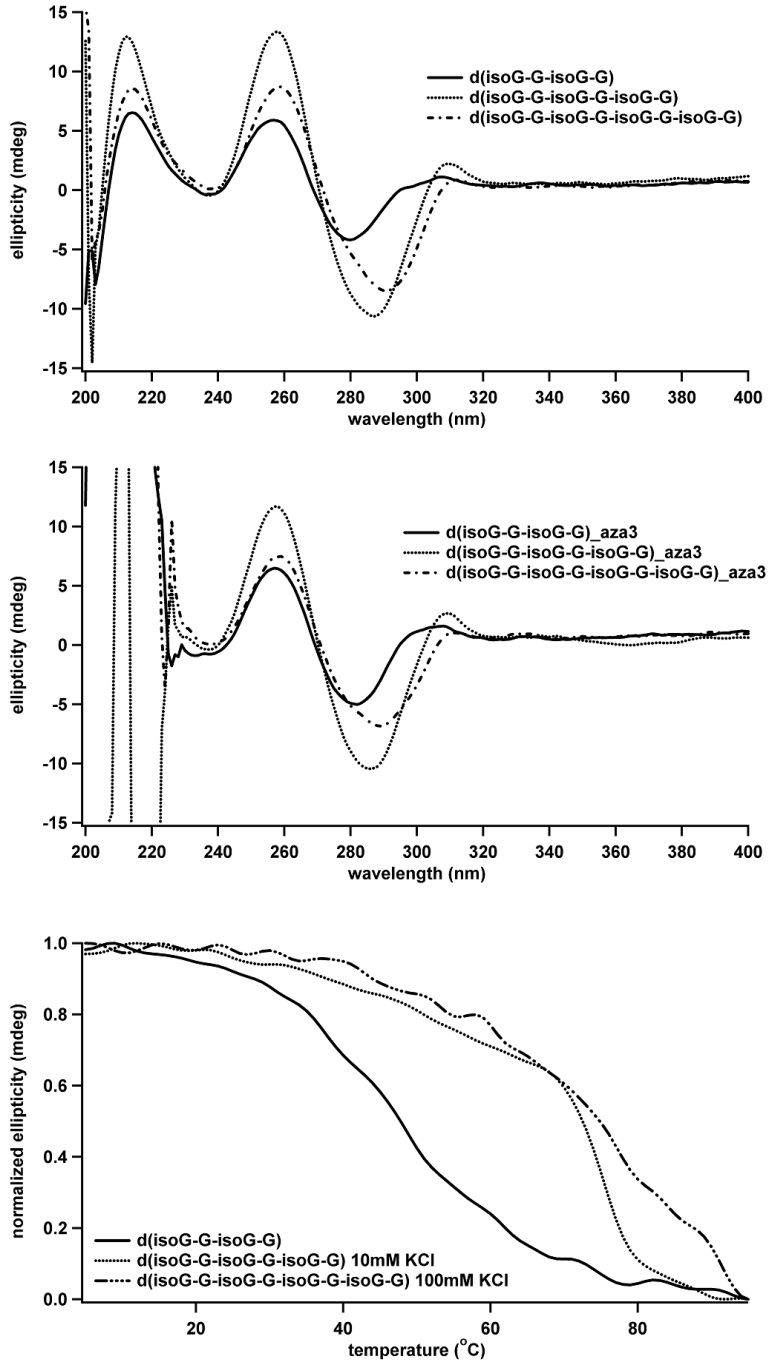

Figure 6.9: CD spectra of isoG·G oligonucleotides in the absence (upper panel) and the presence of aza3 (middle panel) at 5 °C. The T_m curves of d(isoG-G-isoG-G) with various salt concentrations. Samples are all in 60 µM in base in 1X LiCacodylic buffer with 100 mM LiCl. Also 15 µM aza3 is added to the samples with small molecules.

6.5. FUTURE DIRECTIONS

Duplexes formed from all purine bases will bring new facts about understanding molecular recognition as well as the evolution of our genetic system. Even though the purine-purine base pairing recognized as early as in late 1950's, the role and the function of these structures attracted attention in the last decade [60, 105, 155]. Eschenmoser demonstrated the formation of different types of purine-purine base pairing arrangements in homo-DNA oligonucleotides. Just recently, Battersby demonstrated that purine-purine base pairs can also form a duplex structure in natural DNA. Understanding the base pairing in these systems may help us to understand the evolution of A·T and G·C base pairs, why these base pairs are selected during evolution. In addition, identification of new base-pairing arrangements will also help us to extend the genetic alphabet as well as extending the possible use of different nucleic acid structures in creating molecular assemblies for future applications (i.e. nanotechnology)

On the other side, the interactions of small molecules with nucleic acids have always attracted attention due to their role as therapeutic agents. Selective molecular recognition is a key on the way to the rational design of these therapeutic reagents. In addition, designing small molecules which can act as molecular switches will always find its use in growing biotechnology. We are also interested in the use of small molecules in proto-RNA formation. Small molecule interactions with purine-purine will be a good model to investigate the "midwife" hypothesis further.

In this dissertation, we have presented evidence for the formation of a non-canonical base pairing A·A in the presence of coralyne. Coralyne was able to assemble short homod(A) oligonucleotides into anti-parallel polymeric duplexes at neutral pH.

The suggested base pairing for A·A is hold together by two H-bonds between the N7-N6amino and N1-N6amino. We have demonstrated that the same base pairing can also be recognized by other small molecules such as azacyanines. However, the structure (functional groups on small molecule) and size of the small molecule is extremely important in this molecular recognition. In addition the same small molecules can also recognizer other purine-purine base pairings, i.e. aza3 binding to G-quadruplex.

Our preliminary results complimentary to other studies [145] have proven the existence of other possible purine-purine base pairing arrangements in deoxyribo-oligonucleotides. Among those while A·I duplexes revealed very low stability, G·isoG was forming quadruplex structures with very high T_m. Studies with other purine-purine base pairings and their interactions are under investigation in our lab. For instance, D·X (diaminopurine·xanthine) is an ideal candidate base pair to fill the gap between A·I and G·isoG. It is supposed to be more stable than A·I due to the extra H-bond, and should not be able to form higher order structures. It is in our interest first to understand the fundamental reasons under purine-purine base pairings: i.e. under what conditions do they exist, how stable they are, what determines which face of the base will be used for base pairing, the role of the functional groups in different base pairings or the role of the backbone. Second, we are also interested in the recognition of these purine-purine base pairings by small molecules. We certainly interested in investigating the thermodynamic profile of these small molecule-homopurine complexes. Furthermore, we are also interested in synthesis and design of small molecules which are bispurine analogs for the recognition of these purine-purine base pairings. The ultimate goal is to be able to find a "midwife" which will drive the robust assembly of purine bases/mononucleotides into

columnar stacks in solution for proto-RNA formation. The results presented in here are only the preliminary results for reaching this ultimate goal and these projects will be carried out by future students in our lab.

REFERENCES

1. Avery, O.T., C. MacLeod, and M. McCarty, *Studies on the chemical nature of the substance inducing transformation of Pneumococcal types: Induction of transformation by a desoxyribonucleic acd fraction isolated from Pneumococcus type III* Journal of Experimental Medicine, 1944. **79**: p. 137-158.

2. Chargaff, E., *Chemical specifity of nucleic acids and mechanism of their enzymatic degradation.* Experientia, 1950. **6**(6): p. 201-209.

3. Chargaff, E., E. Vischer, R. Doniger, C. Green, and F. Misani, *The composition of desoxypentose nucleic acids of Tymus and Spleen.* Journal of Biological Chemistry, 1949. **177**(1): p. 405-416.

4. Watson, J.D. and F.H.C. Crick, *Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-738.

5. Sines, C.C., L. McFail-Isom, S.B. Howerton, D. VanDerveer, and L.D. Williams, *Cations mediate B-DNA conformational heterogeneity.* Journal of the American Chemical Society, 2000. **122**(45): p. 11048-11056.

6. Voet, D. and J.G. Voet, *Biochemistry: Biomolecules, mechanism of enzyme action, and metabolism*. Vol. 1. 2004: John Wiley & Sons. 1178.

7. Saenger, W., *Principles of Nucleic Acid Structure*. Springer Advanced Texts in Chemistry, ed. C.R. Cantor. 1984, New York: Springer-Verlag. 556.

8. Guittat, L., P. Alberti, D. Gomez, A. De Cian, G. Pennarun, T. Lemarteleur, C. Belmokhtar, R. Paterski, H. Morjani, C. Trentesaux, E. Mandine, F. Boussin, P. Mailliet, L. Lacroix, J.F. Riou, and J.L. Mergny, *Targeting human telomerase for cancer therapeutics.* Cytotechnology, 2004. **45**(1-2): p. 75-90.

9. Keniry, M.A., *Quadruplex structures in nucleic acids.* Biopolymers, 2000. **56**(3): p. 123-146.

10. Rezler, E.M., D.J. Bearss, and L.H. Hurley, *Telomeres and telomerases as drug targets.* Current Opinion in Pharmacology, 2002. **2**(4): p. 415-423.

11. Rezler, E.M., D.J. Bearss, and L.H. Hurley, *Telomere inhibition and telomere disruption as processes for drug targeting.* Annual Review of Pharmacology and Toxicology, 2003. **43**: p. 359-379.

12. Satyanarayana, A., M.P. Manns, and K.L. Rudolph, *Telomeres, telomerase and cancer - An endless search to target the ends.* Cell Cycle, 2004. **3**(9): p. 1138-1150.

13. Roy, A., S. Panigrahi, M. Bhattacharyya, and D. Bhattacharyya, *Structure, stability, and dynamics of canonical and noncanonical base pairs: Quantum chemical studies.* Journal of Physical Chemistry B, 2008. **112**(12): p. 3786-3796.

14. Batey, R.T., R.P. Rambo, and J.A. Doudna, *Tertiary motifs in RNA structure and folding.* Angewandte Chemie-International Edition, 1999. **38**(16): p. 2327-2343.

15. Chandrasekhar, K. and R. Malathi, *Non-Watson Crick base pairs might stabilize RNA structural motifs in ribozymes - A comparative study of group-I intron structures.* Journal of Biosciences, 2003. **28**(5): p. 547-555.

16. Xu, D., T. Landon, N.L. Greenbaum, and M.O. Fenley, *The electrostatic characteristics of G center dot U wobble base pairs.* Nucleic Acids Research, 2007. **35**(11): p. 3836-3847.

17. Waring, M.J. and L.P.G. Wakelin, *Forty Years On,* in *Small Molecule DNA and RNA binders*, M. Demeunynck, C. Bailly, and W.D. Wilson, Editors. 2003, Wiley-VCH Verlag GmbH & Co: Berlin. p. 1-17.

18. Martinez, R. and L. Chacon-Garcia, *The search of DNA-intercalators as antitumoral drugs: What it worked and what did not work.* Current Medicinal Chemistry, 2005. **12**(2): p. 127-151.

19. Thurston, D.E., *Nucleic acid targeting: therapeutic strategies for the 21st century.* British Journal of Cancer, 1999. **80**: p. 65-85.

20. Goodman, L.S., M.M. Wintrobe, W. Dameshek, M.J. Goodman, and A. Gilman, *Nitrogen mustard theraphy: Use of methyl-bis(beta-chloroethyl)amine hydrochloride and tris(beta-chloroethyl)amine hydrochloride for hodgkins disease, lymphosarcoma, leukemia and certain allied and miscellaneous disorders.* Jama-Journal of the American Medical Association, 1946. **132**(3): p. 126-132.

21. Yamori, T., A. Matsunaga, S. Sato, K. Yamazaki, A. Komi, K. Ishizu, I. Mita, H. Edatsugi, Y. Matsuba, K. Takezawa, O. Nakanishi, H. Kohno, Y. Nakajima, H. Komatsu, T. Andoh, and T. Tsuruo, *Potent antitumor activity of MS-247, a novel DNA minor groove binder, evaluated by an in vitro and in vivo human cancer cell line panel.* Cancer Research, 1999. **59**(16): p. 4042-4049.

22. Baraldi, P.G., R. Romagnoli, I. Beria, P. Cozzi, C. Geroni, N. Mongelli, N. Bianchi, C. Mischiati, and R. Gambari, *Synthesis and antitumor activity of new benzoheterocyclic derivatives of distamycin A.* Journal of Medicinal Chemistry, 2000. **43**(14): p. 2675-2684.

23. Mazerska, Z., J. Dziegielewski, and J. Konopa, *Enzymatic activation of a new antitumour drug, 5-diethylaminoethylamino-8-hydroxyimidazoacridinone, C-1311, observed after its intercalation into DNA.* Biochemical Pharmacology, 2001. **61**(6): p. 685-694.

24. Mazerska, Z., P. Sowinski, and J. Konopa, *Molecular mechanism of the enzymatic oxidation investigated for imidazoacridinone antitumor drug, C-1311.* Biochemical Pharmacology, 2003. **66**(9): p. 1727-1736.

25. Lockhart, A.C., M. Howard, K.R. Hande, B.J. Roth, J.D. Berlin, F. Vreeland, A. Campbell, E. Fontana, F. Fiorentini, C. Fowst, V.A. Paty, O. Lankford, and M.L. Rothenberg, *A phase I dose-escalation and pharmacokinetic study of brostallicin (PNU-166196A), a novel DNA minor groove binder, in adult patients with advanced solid tumors.* Clinical Cancer Research, 2004. **10**(2): p. 468-475.

26. Kong, D.H., E.J. Park, A.G. Stephen, M. Calvani, J.H. Cardellina, A. Monks, R.J. Fisher, R.H. Shoemaker, and G. Melillo, *Echinomycin, a small-molecule inhibitor of hypoxia-inducible factor-1 DNA-binding activity.* Cancer Research, 2005. **65**(19): p. 9047-9055.

27. Papadopoulos, K., A. Mita, A. Ricart, D. Hufnagel, D. Northfelt, D. Von Hoff, L. Darjania, J. Lim, C. Padgett, and R. Marschke, *Pharmacokinetic findings from the phase I study of Quarfloxin (CX-3543): A protein-rDNA quadruplex inhibitor, in patients with advanced solid tumors.* Molecular Cancer Therapeutics, 2007. **6**(12): p. 3462S-3463S.

28. Chaires, J.B., *Structural selectivity of drug-nucleic acid interactions probed by competition dialysis*, in *DNA Binders and Related Subjects*. 2005, Springer-Verlag Berlin: Berlin. p. 33-53.

29. Chaires, J.B., *Allostery: DNA does it, too.* Acs Chemical Biology, 2008. **3**(4): p. 207-209.

30. Ren, J.S., X.G. Qu, N. Dattagupta, and J.B. Chaires, *Molecular recognition of a RNA : DNA hybrid structure.* Journal of the American Chemical Society, 2001. **123**(27): p. 6742-6743.

31. Qu, X.G., J.O. Trent, I. Fokt, W. Priebe, and J.B. Chaires, *Allosteric, chiral-selective drug binding to DNA.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(22): p. 12032-12037.

32. Lerman, L.S., *Structural considerations in interaction of DNA and acridines.* Journal of Molecular Biology, 1961. **3**(1): p. 18-30.

33. Neville, D.M. and D.R. Davies, *Interaction of acridine dyes with DNA: An X-Ray diffraction and optical investigation.* Journal of Molecular Biology, 1966. **17**(1): p. 57-74.

34. Wartell, R.M., J.E. Larson, and R.D. Wells, *Netropsin: Specific probe for A-T regions of duplex deoxyribonucleic acid.* Journal of Biological Chemistry, 1974. **249**(21): p. 6719-6731.

35. Strekowski, L. and B. Wilson, *Noncovalent interactions with DNA: An overview.* Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis, 2007. **623**(1-2): p. 3-13.

36. Bloomfield, V.A., D.M. Crothers, and I. Tinoco, *Nucleic Acids: Structures, properties, and functions.* 2000, Sausalito: University Science Books. 794.

37. Jain, S.S., M. Polak, and N.V. Hud, *Controlling nucleic acid secondary structure by intercalation: effects of DNA strand length on coralyne-driven duplex disproportionation.* Nucleic Acids Research, 2003. **31**(15): p. 4608-4615.

38. Dervan, P.B., A.T. Poulin-Kerstien, E.J. Fechter, and B.S. Edelson, *Regulation of gene expression by synthetic DNA-binding ligands,* in *DNA Binders and Related Subjects.* 2005. p. 1-31.

39. Komeda, S., T. Moulaei, K.K. Woods, M. Chikuma, N.P. Farrell, and L.D. Williams, *A third mode of DNA binding: Phosphate clamps by a polynuclear*

platinum complex. Journal of the American Chemical Society, 2006. **128**(50): p. 16092-16103.

40. Hays, F.A., A. Teegarden, Z.J.R. Jones, M. Harms, D. Raup, J. Watson, E. Cavaliere, and P.S. Ho, *How sequence defines structure: A crystallographic map of DNA structure and conformation.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(20): p. 7157-7162.

41. Tabernero, L., N. Verdaguer, M. Coll, I. Fita, G.A. Vandermarel, J.H. Vanboom, A. Rich, and J. Aymami, *Moleular structure of the A-tract DNA dodecamer d(CGCAAATTTGCG) complexed with the minor groove binding drug netropsin.* Biochemistry, 1993. **32**(33): p. 8403-8410.

42. Haq, I., *Part II: The thermodynamics of drug-bipolymer interaction - Thermodynamics of drug-DNA interactions.* Archives of Biochemistry and Biophysics, 2002. **403**(1): p. 1-15.

43. Ihmels, H., K. Faulhaber, D. Vedaldi, F. Dall'Acqua, and G. Viola, *Intercalation of oganic dye molecules into double-stranded DNA. Part 2: The annelated quinolizinium ion as a structural motif in DNA intercalators.* Photochemistry and Photobiology, 2005. **81**(5): p. 1107-1115.

44. Ihmels, H. and D. Otto, *Intercalation of organic dye molecules into double-stranded DNA general principles and recent developments*, in *Supermolecular Dye Chemistry*. 2005, Springer-Verlag Berlin: Berlin. p. 161-204.

45. Crothers, D.M., *Calculation of Binding isotherms for heterogeneous polymers.* Biopolymers, 1968. **6**(4): p. 575-584.

46. Hud, N.V. and F.A.L. Anet, *Intercalation-mediated synthesis and replication: A new approach to the origin of life.* Journal of Theoretical Biology, 2000. **205**(4): p. 543-562.

47. Tsai, C.C., S.C. Jain, and H.M. Sobell, *X-ray crystallographic visualization of drug nucleic acid intercalative binding: Structure of an ethidium-dinucleoside monophosphate crystalline complex, ethidium-5-iodouridylyl(3'-5')adenosine.* Proceedings of the National Academy of Sciences of the United States of America, 1975. **72**(2): p. 628-632.

48. Hud, N.V., S.S. Jain, X.H. Li, and D.G. Lynn, *Addressing the problems of base pairing and strand cyclization in template-directed synthesis - A case for the utility and necessity of 'molecular midwives' and reversible backbone linkages for the origin of proto-RNA.* Chemistry & Biodiversity, 2007. **4**(4): p. 768-783.

49. Jain, S.S., F.A.L. Anet, C.J. Stahle, and N.V. Hud, *Enzymatic behavior by intercalating molecules in a template-directed ligation reaction.* Angewandte Chemie-International Edition, 2004. **43**(15): p. 2004-2008.

50. Seeman, N.C., *Nucleic-acid junctions and lattices.* Journal of Theoretical Biology, 1982. **99**(2): p. 237-247.

51. Zhang, Y.W. and N.C. Seeman, *Construction of a DNA-truncated octahedron.* Journal of the American Chemical Society, 1994. **116**(5): p. 1661-1669.

52. Rothemund, P.W.K., *Folding DNA to create nanoscale shapes and patterns.* Nature, 2006. **440**(7082): p. 297-302.

53. Gothelf, K.V. and T.H. LaBean, *DNA-programmed assembly of nanostructures.* Organic & Biomolecular Chemistry, 2005. **3**(22): p. 4023-4037.

54. Seeman, N.C., *From genes to machines: DNA nanomechanical devices.* Trends in Biochemical Sciences, 2005. **30**(3): p. 119-125.

55. Benvin, A.L., Y. Creeger, G.W. Fisher, B. Ballou, A.S. Waggoner, and B.A. Armitage, *Fluorescent DNA nanotags: Supramolecular fluorescent labels based on intercalating dye arrays assembled on nanostructured DNA templates.* Journal of the American Chemical Society, 2007. **129**(7): p. 2025-2034.

56. Mao, C.D., W.Q. Sun, Z.Y. Shen, and N.C. Seeman, *A nanomechanical device based on the B-Z transition of DNA.* Nature, 1999. **397**(6715): p. 144-146.

57. Yurke, B., A.J. Turberfield, A.P. Mills, F.C. Simmel, and J.L. Neumann, *A DNA-fuelled molecular machine made of DNA.* Nature, 2000. **406**(6796): p. 605-608.

58. Siddiqui-Jain, A., C.L. Grand, D.J. Bearss, and L.H. Hurley, *Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(18): p. 11593-11598.

59. Polak, M. and N.V. Hud, *Complete disproportionation of duplex poly(dT)center dot poly(dA) into triplex poly(dT)center dot poly(dA)center dot poly(dT) and poly(dA) by coralyne.* Nucleic Acids Research, 2002. **30**(4): p. 983-992.

60. Rich, A., F.H.C. Crick, J.D. Watson, and D.R. Davies, *Molecular structure of polyadenylic acid.* Journal of Molecular Biology, 1961. **3**(1): p. 71-86.

61. Ts'o, P.O.P., *Bases, Nucleosides, and Nucleotides*, in *Basic Principles in Nucleic Acid Chemistry*, P.O.P. Ts'o, Editor. 1974, New York: Academic Press. p. 453-584.

62. Ren, J.S. and J.B. Chaires, *Sequence and structural selectivity of nucleic acid binding ligands.* Biochemistry, 1999. **38**(49): p. 16067-16075.

63. Chaires, J.B., *Allosteric conversion of Z-DNA to an intercalated tigh-handed conformation by daunomycin.* Journal of Biological Chemistry, 1986. **261**(19): p. 8899-8907.

64. Pohl, F., T.M. Jovin, W. Baehr, and J.J. Holbrook, *Ethidium bromide as a cooperative effector of a DNA structure.* Proceedings of the National Academy of Sciences of the United States of America, 1972. **69**(12): p. 3805-3809.

65. Persil, O., C.T. Santai, S.S. Jain, and N.V. Hud, *Assembly of an antiparallel homo-adenine DNA duplex by small-molecule binding.* Journal of the American Chemical Society, 2004. **126**(28): p. 8644-8645.

66. Leontis, N.B., J. Stombaugh, and E. Westhof, *The non-Watson-Crick base pairs and their associated isostericity matrices.* Nucleic Acids Research, 2002. **30**(16): p. 3497-3531.

67. Nagaswamy, U., M. Larios-Sanz, J. Hury, S. Collins, Z.D. Zhang, Q. Zhao, and G.E. Fox, *NCIR: a database of non-canonical interactions in known RNA structures.* Nucleic Acids Research, 2002. **30**(1): p. 395-397.

68. Nagaswamy, U., N. Voss, Z.D. Zhang, and G.E. Fox, *Database of non-canonical base pairs found in known RNA structures.* Nucleic Acids Research, 2000. **28**(1): p. 375-376.

69. Hurley, L.H., *DNA and associated targets for drug design.* Journal of Medicinal Chemistry, 1989. **32**(9): p. 2027-2033.

70. Mergny, J.L. and C. Helene, *G-quadruplex DNA: A target for drug design.* Nature Medicine, 1998. **4**(12): p. 1366-1367.

71. Palchaudhuri, R. and P.J. Hergenrother, *DNA as a target for anticancer compounds: methods to determine the mode of binding and the mechanism of action.* Current Opinion in Biotechnology, 2007. **18**(6): p. 497-503.

72. Park, S.H., P. Yin, Y. Liu, J.H. Reif, T.H. LaBean, and H. Yan, *Programmable DNA self-assemblies for nanoscale organization of ligands and proteins.* Nano Letters, 2005. **5**(4): p. 729-733.

73. Persil, O. and N.V. Hud, *Harnessing DNA intercalation.* Trends in Biotechnology, 2007. **25**(10): p. 433-436.

74. Dower, K., N. Kuperwasser, H. Merrikh, and M. Rosbash, *A synthetic A tail rescues yeast nuclear accumulation of a ribozyme-terminated transcript.* Rna-a Publication of the Rna Society, 2004. **10**(12): p. 1888-1899.

75. Wickens, M., P. Anderson, and R.J. Jackson, *Life and death in the cytoplasm: Messages from the 3' end.* Current Opinion in Genetics & Development, 1997. **7**(2): p. 220-232.

76. Topalian, S.L., M.I. Gonzales, Y. Ward, X.A. Wang, and R.F. Wang, *Revelation of a cryptic major histocompatibility complex class II-restricted tumor epitope in a novel RNA-processing enzyme.* Cancer Research, 2002. **62**(19): p. 5505-5509.

77. Topalian, S.L., S. Kaneko, M.I. Gonzales, G.L. Bond, Y. Ward, and J.L. Manley, *Identification and functional characterization of neo-poly(A) polymerase, an RNA processing enzyme overexpressed in human tumors.* Molecular and Cellular Biology, 2001. **21**(16): p. 5614-5623.

78. Xing, F.F., G.T. Song, J.S. Ren, J.B. Chaires, and X.G. Qu, *Molecular recognition of nucleic acids: Coralyne binds strongly to poly(A).* Febs Letters, 2005. **579**(22): p. 5035-5039.

79. Giri, P., M. Hossain, and G.S. Kumar, *Molecular aspects on the specific interaction of cytotoxic plant alkaloid palmatine to poly(A).* International Journal of Biological Macromolecules, 2006. **39**(4-5): p. 210-221.

80. Giri, P. and G.S. Kumar, *Specific binding and self-structure induction to poly(A) by the cytotoxic plant alkaloid sanguinarine.* Biochimica Et Biophysica Acta-General Subjects, 2007. **1770**(9): p. 1419-1426.

81. Yadav, R.C., G.S. Kumar, K. Bhadra, P. Giri, R. Sinha, S. Pal, and M. Maiti, *Berberine, a strong polyriboadenylic acid binding plant alkaloid: spectroscopic, viscometric, and thermodynamic study.* Bioorganic & Medicinal Chemistry, 2005. **13**(1): p. 165-174.

82. Ben-Shaul, A. and W.M. Gelbart, *Statistical thermodynamics of amphiphile self-assembly: Structure and phase transitions in micellar solutions*, in *Micelles, membranes, microemulsions and monolayers*, W.M. Gelbart, A. Ben-Shaul, and D. Roux, Editors. 1994, Springer-Verlag: New York. p. 1-90.

83. Debye, P., *Light scattering in soap solutions.* Annals of the New York Academy of Sciences, 1949. **51**(4): p. 575-592.

84. Kegel, W.K. and P. van der Schoot, *Competing hydrophobic and screened-Coulomb interactions in hepatitis B virus capsid assembly.* Biophysical Journal, 2004. **86**(6): p. 3905-3913.

85. Tanford, C., *The hydrophobic effect formation of micelles and biological membranes.* 2nd ed. 1980, New York: John Wiley & Sons. 234.

86. MoraruAllen, A.A., S. Cassidy, J.L.A. Alvarez, K.R. Fox, T. Brown, and A.N. Lane, *Coralyne has a preference for intercalation between TA center dot T triples in intramolecular DNA triple helices.* Nucleic Acids Research, 1997. **25**(10): p. 1890-1896.

87. Lee, J.S., L.J.P. Latimer, and K.J. Hampel, *Coralyne binds tightly to both T.A.T containing and C.G.C+ containing DNA triplexes.* Biochemistry, 1993. **32**(21): p. 5591-5597.

88. Biver, T., C. Ciatto, F. Secco, and M. Venturini, *Dye-induced aggregation of single stranded RNA: A mechanistic approach.* Archives of Biochemistry and Biophysics, 2006. **452**(2): p. 93-101.

89. Garbett, N.C. and D.E. Graves, *Extending nature's leads: the anticancer agent ellipticine.* Current Medicinal Chemistry - Anti-cancer Agents, 2004. **4**(2): p. 149-172.

90. Kohn, K.W., M.J. Waring, D. Glaubiger, and C.A. Friedman, *Intercalative binding of ellipticine to DNA.* Cancer Research, 1975. **35**(1): p. 71-76.

91. Canals, A., M. Purciolas, J. Aymani, and M. Coll, *The anticancer agent ellipticine unwinds DNA by intercalative binding in an orientation parallel to base pairs.* Acta Crystallographica Section D-Biological Crystallography, 2005. **61**: p. 1009-1012.

92. Jain, S.C., K.K. Bhandary, and H.M. Sobell, *Visualization of drug-nucleic acid interactions at atomic resolution: Structure of drug-dinucleoside monophosphate crystalline complexes, ellipticine-5-iodocytidylyl (3'-5') guanosine and 3,5,6,8-tetramethyl-n-methyl phenanthrolinium-5-iodocytidylyl (3'-5') guanosine.* Journal of Molecular Biology, 1979. **135**(4): p. 813-840.

93. Muller, W. and D.M. Crothers, *Studies of binding of actinomycin and related compounds to DNA.* Journal of Molecular Biology, 1968. **35**(2): p. 251-290.

94. Galietta, L.J.V., M.F. Springsteel, M. Eda, E.J. Niedzinski, K. By, M.J. Haddadin, M.J. Kurth, M.H. Nantz, and A.S. Verkman, *Novel CFTR chloride channel activators identified by screening of combinatorial libraries based on flavone and benzoquinolizinium lead compounds.* Journal of Biological Chemistry, 2001. **276**(23): p. 19723-19728.

95. Haddadin, M.J., J.M. Kurth, and M.M. Olmstead, *One-step synthesis of new heterocyclic azacyanines.* Tetrahedron Letters, 2000. **41**(30): p. 5613-5616.

96. Huang, K.S., M.J. Haddadin, M.M. Olmstead, and M.J. Kurth, *Synthesis and reactions of some heterocyclic azacyanines.* Journal of Organic Chemistry, 2001. **66**(4): p. 1310-1315.

97. Maggini, R., Secco, M. Venturini, and H. Diebler, *Kinetic-study of double helix formation and double helix dissociation of polyadenylic acid.* Journal of the Chemical Society-Faraday Transactions, 1994. **90**(16): p. 2359-2363.

98. Lewis, K. and F.M. Ausubel, *Prospects for plant-derived antibacterials.* Nature Biotechnology, 2006. **24**(12): p. 1504-1507.

99. Ivanovska, N. and S. Philipov, *Study on the anti-inflammatory action of Berberis vulgaris root extract, alkaloid fractions and pure alkaloids.* International Journal of Immunopharmacology, 1996. **18**(10): p. 553-561.

100. Kuo, C.L., C.C. Chou, and B.Y.M. Yung, *Berberine complexes with DNA in the berberine induced apoptosis in human leukenic HL-60 cells.* Cancer Letters, 1995. **93**(2): p. 193-200.

101. Wu, H.L., C.Y. Hsu, W.H. Liu, and B.Y.M. Yung, *Berberine-induced apoptosis of human leukemia HL-60 cells is associated with down-regulation of nucleophosmin/B23 and telomerase activity.* International Journal of Cancer, 1999. **81**(6): p. 923-929.

102. Persil Cetinkol, O., A.E. Engelhart, R.K. Nanjunda, W.D. Wilson, and N.V. Hud, *Submicromolar, selective G-quadruplex ligands from one pot: Thermodynamic and structural studies of human telomeric DNA binding by azacyanines.* Chembiochem, 2008 (in publication).

103. Qu, X.G. and J.B. Chaires, *Analysis of drug-DNA binding data*, in *Numerical Computer Methods, Part C*. 2000, Academic Press Inc: San Diego. p. 353-369.

104. Arnott, S., Chandras.R, and C.M. Marttila, *Structures of polyinosinic acid and polyguanylic acid.* Biochemical Journal, 1974. **141**(2): p. 537-543.

105. Gellert, M., M.N. Lipsett, and D.R. Davies, *Helix formation by guanylic acid* Proceedings of the National Academy of Sciences of the United States of America, 1962. **48**(12): p. 2013-2018.

106. Zimmerman, S.B., G.H. Cohen, and D.R. Davies, *X-ray fiber diffraction and model building study of polyguanylic acid and polyinosinic acid.* Journal of Molecular Biology, 1975. **92**(2): p. 181-192.

107. Burge, S., G.N. Parkinson, P. Hazel, A.K. Todd, and S. Neidle, *Quadruplex DNA: sequence, topology and structure.* Nucleic Acids Research, 2006. **34**(19): p. 5402-5415.

108. Davis, J.T., *G-quartets 40 years later: From 5 '-GMP to molecular biology and supramolecular chemistry.* Angewandte Chemie-International Edition, 2004. **43**(6): p. 668-698.

109. Laughlan, G., A.I.H. Murchie, D.G. Norman, M.H. Moore, P.C.E. Moody, D.M.J. Lilley, and B. Luisi, *The high resolution crystal structure of a parallel stranded guanine tetraplex.* Science, 1994. **265**(5171): p. 520-524.

110. Neidle, S. and G.N. Parkinson, *The structure of telomeric DNA.* Current Opinion in Structural Biology, 2003. **13**(3): p. 275-283.

111. Smith, F.W. and J. Feigon, *Quadruplex structure of Oxytricha telomeric DNA oligonucleotides.* Nature, 1992. **356**(6365): p. 164-168.

112. Hud, N.V. and J. Plavec, eds. *The role of cations in determining quadruplex structure and stability*. Quadruplex Nucleic Acids, ed. S. Neidle and S. Balasubramanian. 2006, The Royal Society of Chemistry: Cambridge. 100-130.

113. Huppert, J.L., *Four-stranded DNA: cancer, gene regulation and drug development.* Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences, 2007. **365**(1861): p. 2969-2984.

114. Moyzis, R.K., J.M. Buckingham, L.S. Cram, M. Dani, L.L. Deaven, M.D. Jones, J. Meyne, R.L. Ratliff, and J.R. Wu, *A highly conserved repetitive DNA sequence, (TTAGGG)N, present at the telomeres of human chromosomes* Proceedings of the National Academy of Sciences of the United States of America, 1988. **85**(18): p. 6622-6626.

115. Kim, N.W., M.A. Piatyszek, K.R. Prowse, C.B. Harley, M.D. West, P.L.C. Ho, G.M. Coviello, W.E. Wright, S.L. Weinrich, and J.W. Shay, *Specific association of human telomerase activity with immortal cells and cancer.* Science, 1994. **266**(5193): p. 2011-2015.

116. Oganesian, L., I.K. Moon, T.M. Bryan, and M.B. Jarstfer, *Extension of G-quadruplex DNA by ciliate telomerase.* Embo Journal, 2006. **25**(5): p. 1148-1159.

117. Zahler, A.M., J.R. Williamson, T.R. Cech, and D.M. Prescott, *Inhibition of telomerase by G-quartet DNA structures.* Nature, 1991. **350**(6320): p. 718-720.

118. Zaug, A.J., E.R. Podell, and T.R. Cech, *Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(31): p. 10864-10869.

119. De Cian, A., G. Cristofari, P. Reichenbach, E. De Lemos, D. Monchaud, M.P. Teulade-Fichou, K. Shin-Ya, L. Lacroix, J. Lingner, and J.L. Mergny, *Reevaluation of telomerase inhibition by quadruplex ligands and their mechanisms of action.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(44): p. 17347-17352.

120. Han, F.X.G., R.T. Wheelhouse, and L.H. Hurley, *Interactions of TMPyP4 and TMPyP2 with quadruplex DNA. Structural basis for the differential effects on telomerase inhibition.* Journal of the American Chemical Society, 1999. **121**(15): p. 3561-3570.

121. Izbicka, E., R.T. Wheelhouse, E. Raymond, K.K. Davidson, R.A. Lawrence, D.Y. Sun, B.E. Windle, L.H. Hurley, and D.D. Von Hoff, *Effects of cationic porphyrins as G-quadruplex interactive agents in human tumor cells.* Cancer Research, 1999. **59**(3): p. 639-644.

122. Harrison, R.J., A.P. Reszka, S.M. Haider, B. Romagnoli, J. Morrell, M.A. Read, S.M. Gowan, C.M. Incles, L.R. Kelland, and S. Neidle, *Evaluation of by disubstituted acridone derivatives as telomerase inhibitors: the importance of G-quadruplex binding.* Bioorganic & Medicinal Chemistry Letters, 2004. **14**(23): p. 5845-5849.

123. Koeppel, F., J.F. Riou, A. Laoui, P. Mailliet, P.B. Arimondo, D. Labit, O. Petitgenet, C. Helene, and J.L. Mergny, *Ethidium derivatives bind to G-quartets, inhibit telomerase and act as fluorescent probes for quadruplexes.* Nucleic Acids Research, 2001. **29**(5): p. 1087-1096.

124. Dixon, I.M., F. Lopez, A.M. Tejera, J.P. Esteve, M.A. Blasco, G. Pratviel, and B. Meunier, *A G-quadruplex ligand with 10000-fold selectivity over duplex DNA.* Journal of the American Chemical Society, 2007. **129**(6): p. 1502-1503.

125. Bugaut, A., K. Jantos, J.L. Wietor, R. Rodriguez, J.K.M. Sanders, and S. Balasubramanian, *Exploring the differential recognition of DNA G-quadruplex targets by small molecules using dynamic combinatorial chemistry.* Angewandte Chemie-International Edition, 2008. **47**(14): p. 2677-2680.

126. Horowitz, E.D. and N.V. Hud, *Ethidium and proflavine binding to a 2 ',5 '-linked RNA duplex.* Journal of the American Chemical Society, 2006. **128**(48): p. 15380-15381.

127. Goddard, T.D. and D.G. Kneller, *SPARKY*. 2007: University of California, San Francisco.

128. Marat, K., *Spin Works*. 2002: University of Manitoba, Canada.

129. Luu, K.N., A.T. Phan, V. Kuryavyi, L. Lacroix, and D.J. Patel, *Structure of the human telomere in K+ solution: An intramolecular (3+1) G-quadruplex scaffold.* Journal of the American Chemical Society, 2006. **128**(30): p. 9963-9970.

130. Tanious, F.A., D. Hamelberg, C. Bailly, A. Czarny, D.W. Boykin, and W.D. Wilson, *DNA sequence dependent monomer-dimer binding modulation of asymmetric benzimidazole derivatives.* Journal of the American Chemical Society, 2004. **126**(1): p. 143-153.

131. Freyer, M.W., R. Buscaglia, K. Kaplan, D. Cashman, L.H. Hurley, and E.A. Lewis, *Biophysical studies of the c-MYC NHE III1 promoter: Model quadruplex interactions with a cationic porphyrin.* Biophysical Journal, 2007. **92**(6): p. 2007-2015.

132. Dai, J.X., C. Punchihewa, A. Ambrus, D. Chen, R.A. Jones, and D.Z. Yang, *Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation.* Nucleic Acids Research, 2007. **35**(7): p. 2440-2450.

133. Duan, W.H., A. Rangan, H. Vankayalapati, M.Y. Kim, Q.P. Zeng, D.K. Sun, H.Y. Han, O.Y. Fedoroff, D. Nishioka, S.Y. Rha, E. Izbicka, D.D. Von Hoff, and L.H. Hurley, *Design and synthesis of fluoroquinophenoxazines that interact with human telomeric G-quadruplexes and their biological effects.* Molecular Cancer Therapeutics, 2001. **1**(2): p. 103-120.

134. Kim, M.Y., Y. Na, H. Vankayalapati, M. Gleason-Guzman, and L.H. Hurley, *Design, synthesis, and evaluation of psorospermin/quinobenzoxazine hybrids as structurally novel antitumor agents.* Journal of Medicinal Chemistry, 2003. **46**(14): p. 2958-2972.

135. Ts'o, P., I. Melvin, and A. Olson, *Interaction and association of bases and nucleosides in aqueous solutions.* J. Am. Chem. Soc., 1963. **85**: p. 1289-1296.

136. Crick, F.H.C., *The origin of the genetic code.* J. Mol. Biol., 1968. **38**: p. 367-379.

137. Orgel, L.E., *Evolution of the genetic apparatus.* J. Mol. Biol., 1968. **38**: p. 381-393.

138. Wächtershäuser, G., *An all-purine precursor of nucleic acids.* Proc. Natl. Acad. Sci. USA, 1988. **84**: p. 1134-1135.

139. Zubay, G. and T. Mui, *Prebiotic synthesis of nucleosides.* Origins Life Evol. B., 2001. **31**: p. 87-102.

140. Battersby, T.R., M. Albalos, and M.J. Friesenhahn, *An unusual mode of DNA duplex association: Watson-Crick interaction of all-purine deoxyribonucleic acids.* Chem. Biol., 2007. **14**: p. 525–531.

141. Rich, A., *Formation of two- and three-stranded helical molecules by polyinosinic acid and polyadenylic acid.* Nature, 1958. **181**: p. 521-525.

142. Howard, F.B. and H.T. Miles, *Interaction of poly(A) and poly(I): Reinvestigation.* Biochemistry, 1977. **16**(21): p. 4647-4650.

143. Groebke, K., J. Hunziker, W. Fraser, L. Peng, U. Diederichsen, K. Zimmermann, A. Holzner, C. Leumann, and A. Eschenmoser, *Why pentose- and not hexose-nucleic acids? Purine-purine pairing in homo-DNA: guanine, isoguanine, 2,6-diaminopurine, and xanthine.* Helv. Chim. Acta, 1998. **81**: p. 375-474.

144. Eschenmoser, A. and E. Loewenthal, *Chemistry of potentially prebiological natural-products* Chemical Society Reviews, 1992. **21**(1): p. 1-16.

145. Battersby, T.R., M. Albalos, and M.J. Friesenhahn, *An unusual mode of DNA duplex association: Watson-Crick interaction of all-purine deoxyribonucleic acids.* Chemistry & Biology, 2007. **14**(5): p. 525-531.

146. Corfield, P.W.R., W.N. Hunter, T. Brown, P. Robinson, and O. Kennard, *Inosine-adenine base pairs in a B-DNA duplex.* Nucleic Acids Research, 1987. **15**(19): p. 7935-7949.

147. Uesugi, S., Y. Oda, M. Ikehara, Y. Kawase, and E. Ohtsuka, *Identification of I-A mismatch base pairing structure in DNA.* Journal of Biological Chemistry, 1987. **262**(15): p. 6965-6968.

148. Maki, H. and A. Kornberg, *The polymerase subunit of DNA Polymerase-III of Escherichia Coli: Purification of the alpha-subunit, devoid of nuclease activities.* Journal of Biological Chemistry, 1985. **260**(24): p. 2987-2992.

149. Seela, F., Y. He, and C.F. Wei, *Parallel-stranded oligonucleotide duplexes containing 5-methylisocytosine-guanine and isoguanine-cytosine base pairs.* Tetrahedron, 1999. **55**(31): p. 9481-9500.

150. Gu, J.D. and J. Leszczynski, *Isoguanine: From base pair to tetrad.* Journal of Physical Chemistry A, 2003. **107**(44): p. 9447-9455.

151. Gu, J.D. and J. Leszczynski, *Isoguanine complexes: Quintet versus tetrad.* Journal of Physical Chemistry B, 2003. **107**(27): p. 6609-6613.

152. Roberts, C., J.C. Chaput, and C. Switzer, *Beyond guanine quartets: cation-induced formation of homogenous and chimeric DNA tetraplexes incorporating iso-guanine and guanine.* Chemistry & Biology, 1997. **4**(12): p. 899-908.

153. Seela, F. and R. Kroschel, *Quadruplex and pentaplex self-assemblies of oligonucleotides containing short runs of 8-aza-7-deaza-2 '-deoxyisoguanosine or 2 '-deoxyisoguanosine.* Bioconjugate Chemistry, 2001. **12**(6): p. 1043-1050.

154. Seela, F., C.F. Wei, and A. Melenewski, *Isoguanine quartets formed by d(T(4)isoG(4)T(4)): Tetraplex identification and stability.* Nucleic Acids Research, 1996. **24**(24): p. 4940-4945.

155. Rich, A., *Formation of 2-stranded and 3-stranded helical molecules by polyinosinic acid and polyadenylic acid.* Nature, 1958. **181**(4608): p. 521-525.